Molecular ecology of the honey bee *Apis mellifera* L. in Ireland

Keith A. Browne BSc (Hons)

Supervisor:

Professor Grace P. McCormack
School of Natural Sciences
Discipline of zoology
National University of Ireland Galway

December 2021

Table of contents — Page

Acknowledgements

My humble and heartfelt thanks go to:

My wife Clare first and foremost, for her endless support, love, encouragement and for putting up with my numerous troughs. Her fortitude enabled me to keep going at what is the hardest endeavour I have ever undertaken. Without her, this adventure would never have even got off the ground.

A close second, Grace, for the supervision (emphasis on 'super') of my work. Her door was always open to dispatch help, suggestions, or moral support. In the background she worked hard finding funding for my far-flung ideas and making collaborations possible. If there is any finesse to this thesis it is due to her influence.

My international collaborators; Prof. M. Alice Pinto and her (then) PhD student Dora Henriques along with everyone in Instituto Politécnico de Bragança, Portugal for making me so welcome during my brief sojourn working on SNPs; Prof. Alex Keller and his (then) PhD student Anna Voulgari-Kokota in Universität Würzburg for their help and hospitality working on the pollen and microbiome metabarcoding. I hope our collaborations continue.

Eoin McLaughlin and Maeve Edwards for keeping the lab supplied with necessaries, help in housing samples, finding things, fixing other things, helping to keep the apiary going, driving around the country for bee samples, the list goes on and on. More importantly, for the coffeetime chats and craic. May the coffee-pot never run dry and there always be an endless supply of (full-fat) milk and apple tart.

All the wonderful lecturers in the Discipline of Zoology who freely offered advice and encouragement whether it was because I knocked on their door enquiring or just in brief pass-by corridor conversations.

My GRC of Charlie Spillane, Mike Geary and especially Mike Gormally for their guidance and advice.

NIHBS, the Eva Crane Trust and the Martin Ryan Institute for helping fund my laptop, laboratory consumables and conference travel.

NIHBS again, and its membership, for the enormous supply of bees, varroa counts and colony records upon which this thesis sits. Particular thanks to Gerard Coyne and Sean Osborne for getting this fledgling beekeeper off the ground.

Caroline Delaney for invaluable assistance by transferring the endless data mentioned above into digital form.

Mary, Suzanne, Colm and all the admin staff in The Ryan Institute for their help with day-to-day admin life and in helping to prep some of my funding applications.

The beekeepers who permitted me to raid their colonies every month for bees, brood, and pollen. Mark Newenham, Pat Deasy, PJ Curran, John Howlin, Seamus Kennedy, John Little, PJ Murphy and Noel Leahy. Chapter 3 wouldn't exist with their trust in me not to kill their stocks.

Aaron Golden for attempting to help me with of my poor grasp of R and producing the diversity index graphs.

Roberto Ferrari for making the work we did together so enjoyable.

Morag Taite for being the best lab colleague you could hope for, ready to pitch in and help with whatever including but not limited to documenting samples, loaning lab consumables and of course, "pulling the arse out of bees".

Tom Rossiter for assisting me with the GIS map making.

The various fourth year students that in return for the little guidance I could give them on their projects extracted DNA and collected free-living bee samples.

Last but by no means least, my "little Italian sister" Maria Vittoria for showing me the vagaries of Qiime and sharing the PhD journey.

Dedication

For my father, who never got to see this happen.

Abstract

The honey bee sub-species native to Ireland is *Apis mellifera mellifera*, referred to locally as the Black bee. It is the same sub-species that has undergone widespread extinction across the rest of its native range in northern Europe as a consequence of habitat loss, replacement, hybridisation, and colony death caused by the novel parasite *Varroa destructor*.

A programme to locate and selectively breed for bees resistant to *V. destructor* was initiated and investigations were made into the genotypes of the monitored colonies. Using a combination of mtDNA, microsatellite and SNP markers, the queen lineage and extent of hybridisation between M and C lineages in a selection of bees was determined.

As a possible genetic source of resistance to the parasite, free-living, unmanaged colonies were located using a citizen science approach. They were also genotyped using the same markers and their survival was monitored.

Eight apiaries were sampled for bees, brood, and pollen to investigate pollen use by Irish honey bees between apiaries and colonies located in similar rural landscapes and to make the first concurrent study of the gut bacteria of these bees.

An extensive trial indicated that the process of creating a widespread breeding programme may be prohibitively difficult to achieve in an Irish context. The reasons behind this are discussed and a suggestion is made for a modular, localised model. The majority of bees genotyped, whether managed or free-living, are revealed as M-lineage by mtDNA genotyping and assigned as the sub-species native to Ireland. The citizen science approach resulted in the discovery of a large free-living bee population living in natural and artificial cavities. The survival of a number of these colonies was monitored for periods exceeding the three year limit expected of colonies infested with *V. destructor*. The pollen data added to our knowledge from the single previous peer-reviewed research on honey bee pollen in Ireland by revealing the use of a significant number of new plant species. These new data permit a temporal comparison between and within apiaries and are compared to a concurrent gut bacteria profile.

General introduction

Honey bee biology

Bees and other hymenoptera such as wasps and ants along with termites, some crustaceans (Duffy 2003) and mammal species such as mole rats, are some of the few animals that have achieved truly social, or eusocial, status in the long history of animals. Despite the small number of insect species that are eusocial it is estimated that they form the largest proportion of insect biomass on earth (Nowak *et al*. 2010). The western honey bee *Apis mellifera* L. is a holometabolous insect of the order Hymenoptera and is one of these few eusocial animals.

Key to the success of eusocial insects, is the production of large numbers of individuals operating in and protecting a perennial nest (Nowak *et al.* 2010). Honey bees are no exception. They are cavity dwellers, living in tree hollows, rock cavities, spaces in artificial structures such as buildings or in specially constructed hives. They make a nest in these spaces by fixing in place vertical, double-sided sheets of hexagonal cells which are created from wax produced by glands located ventrally on their bodies, on the inner sides of the abdominal sternites.

Within the wax cells the storage and maturation of foodstuffs collected from the surrounding environment takes place. They are also used for raising brood from egg to adult. The queen lays an individual egg in each brood cell which develops into a larva, which is fed by nurse worker bees. Prior to the pupation stage of each bee's development, the cell is capped with wax by workers and the prepupae/pupae spins a pupal cocoon within. When fully developed, the adult bee chews through the wax cap and emerges.

While numbers vary with colony health and the volume of the cavity, a healthy hive in an artificial apiary hive of about 40L to 60L capacity, can have between 30,000-50,000 bees. A diplo-haploid organism, the colony consists of a (usually) single diploid queen, tens of thousands of diploid female workers and a few hundred or more haploid male drones. The queen is polyandrous, mating with up to 20 drones during multiple mating flights she makes to drone congregation areas (DCAs). The diplo-haploid and polyandrous nature of honey bees means that the haploid male drone lineage stems directly from the queen's alone while female workers are a mixture of full and half-sisters, although any clear delineation is tempered by their high miotic recombination rate (Beye *et al.* 2006) which serves to create additional genetic diversity and with possible influence on colony fitness (Gadau *et al.* 2000).

Pollination services by the honey bee

As well as producing commercial volumes of honey and other hive products, honey bees are considered the world's foremost pollinator of crops although care in overstating their role (Breeze *et al.* 2011) at the expense of other pollinators is warranted. The basis of their commercial pollination service is the large numbers of individual foragers in each hive, that hive numbers can be increased through queen breeding and that hives can be transported to provide pollination services on specific crop plants such as rapeseed or almonds. When considering a country's need for pollination services, the dependence on the honey bee may be extenuated in countries, such as on the American continents, where it has been introduced for that very purpose. Consequently, that dependence appears to be a consequence of economic, rather than biological, factors (Aizen and Harder 2009) such as a dearth of native pollinators. However, most research output comes from developed countries (Porto *et al.* 2020) and the situation in less-developed countries may be unclear. In Ireland the economic value of the bee pollinator ecosystem service is an estimated €85 million per annum (Harrington 2016) although it is unclear if this includes less tractable values such as the enhancement of the landscape through the promotion of natural floral diversity. This pollinator service is of vital importance to maintain Ireland's recreational spaces and, by extension, tourism, which is one of the country's highest income generators.

Honey bee pathogens

Bees suffer from a significant number of number of pathogens such as the trypanosomatid parasite, *Crithidia mellificae,* endoparasital tracheal mite *Acarapsis woodii,* bacteria in the form of American foulbrood (AFB) *Paenibacillus larvae* subsp. *larvae* (formerly *Bacillus larvae*) and European foulbrood (EFB) *Mellissococcus plutonius.* Nosmoosis is commonplace in bee brood, caused by the microsporidian (fungal) parasites *Nosema apis* and *N. ceranae* as is fungal infection in the form of chalkbrood (*Ascosphaera apis*). A protozoan, *Malpighamoeba mellificae*, causes amoeba disease and there are multiple viral diseases such as deformed wing, acute paralysis, Israeli acute paralysis and sacbrood. For the western honey bee population worldwide, the primary cause of concern is the mite *Varroa destructor.*

A drawback of eusociality is that the density of individuals within the living space results in disease onset passing quickly though the entire nest. Although possessing reduced individual immune responses relative to other insects (Evans *et al.* 2006), the honey bee has developed

group-level immunity through behaviours such as grooming or the removal of infected and sick individuals (Boecking and Spivak 1999). These hygienic behaviours, as they are known, are so highly tuned that even their single queen can be killed, removed, or forced out if illness or even lack of sperm to fertilise her eggs is sensed by the workers. Additionally, greater genetic diversity between nestmates, something which is strongly influenced by how polyandrous the queen has been, can produce healthier colonies (Tarpy 2003). This could occur by limiting the numbers of individuals that succumb to any given pathogen through varied levels of resistance to that pathogen across the colony.

Varroa destructor biology

An obligate ectoparasite mite (Acaridae), *V. destructor* (Anderson and Trueman 2000), was introduced into the *Apis mellifera* population through close contact with the mite's adapted host, the Eastern honey bee *A. cerana.* Its hemimetabolic development takes place entirely within the hive environment with egg laying, nymphal and adult stages along with mating all taking place during the capped cell period of honey bee brood development. In the cell, the mites feed on the tissue of the honey bee prepupae/pupae (Huang 2012) and further parasitic feeding on the fat bodies of adult bees (Ramsey *et al.* 2019) allows the adult female mites to survive whilst they are carried from brood cell to brood cell and between colonies.

During feeding on the larvae and adult bees, varroa acts as a vector for a suite of viruses (Tentcheva *et al.* 2004), some of which are mentioned above, with which it inoculates the host bee. While the parasitic feeding on the bee reduces the host's body mass and lifespan (Dejong *et al.* 1982), the effects of these viruses is likely to have the more serious effect on the health of a colony as these can be spread further through oral trophallaxis by the bees. By this combination of factors, the consequence of varroa infestation can be a significant reduction in colony health followed by death within one to three years.

Major threats to all bees and honey bees

Lack of suitable habitat for nesting and foraging is one of the main threats to the survival of all pollinators including honey bees and is primarily created through agricultural expansion (Nieto *et al*. 2014), often involving the removal of remaining natural habitats (Sydenham *et al.* 2014) such as can occur with hedgerows in Ireland. Insecticide and herbicide use also form part of agricultural expansion and can have lethal and sub-lethal effects on bees including

reducing foraging efficacy and negatively affecting navigational ability (Goulson 2013, Gill and Raine 2014, Cullen *et al.* 2019). In addition to lethal and sub-lethal effects, the significance of the removal of food plants for bees through herbicidal use is still undetermined (Cullen *et al.* 2019).

The movement of honey bees and their commercial products between countries has the capability to facilitate the introduction of novel pathogens into populations. This has already occurred with the introduction of *V. destructor* into the *A. mellifera* population. Currently, following human-mediated introductions, the giant asian hornet *Vespa velutina* and small hive beetle *Aethina tumida* are persisting outside of their natural ranges and form serious threats to naïve honey bee populations in Europe (Cuthbertson *et al.* 2013, Requier *et al.* 2019).

Ruttner (1988) identified 24 sub-species of *A. mellifera* worldwide using morphological analysis and that number is growing through the increasing use of molecular markers (Franck *et al.* 2000, Meixner *et al.* 2011, Alburaki *et al.* 2013, Ilyasov *et al.* 2020).

Unfortunately, the practice of transporting bee sub-species outside of their natural geographical range for breeding purposes has resulted in introgressive hybridisation of local sub-species, potentially causing the loss of local adaptations (De la Rua *et al.* 2009, Meixner *et al.* 2010). *A. m. mellifera*, the sub-species native to Northern Europe, including Ireland, is under threat from the importation of the southern European sub-species *A. m. ligustica* and *A. m. carnica* along with commercial hybrids such as “Buckfast” and “Starline”. Of major cause for concern is that, within the EU, the common agricultural policy (CAP) allows for the free trade of honey bees between member states and thus forms a considerable block to correcting the problem.

Overview of beekeeping in Ireland

The wet Atlantic climate of Ireland is not as conducive to honey production as that of continental Europe, consequently the great majority of beekeepers in Ireland are hobbyists, keeping less than 10 hives on average (Chauzat *et al.* 2013). The few commercial beekeeping operations tend to be in the south and south-east of the island where the climate is dryer and warmer. The demanding climate forms the backdrop for many beekeepers’ assertions about the local adaptations of their bees. The approximately 4500 registered beekeepers and their local beekeeping associations (BKAs) are represented by three umbrella bodies, the Ulster

Beekeeping Association (UBKA), the Federation of Irish Beekeeping Associations (FIBKA) and the Irish Beekeeping Association (IBA). Members are free to choose the sub-species or hybrid honey bee type that they wish to use. Although many BKAs inform their members of the existence of the local sub-species it was still felt by some that the threat to the genetic integrity of *A. m. mellifera* is too great to leave the matter to individual choice. In 2012, the Native Irish Honey Bee Society (NIHBS) was formed to promote the conservation and preservation of *A.m.mellifera* on the island of Ireland and generally consists of beekeepers who are still members of the umbrella groups. The number of unregistered beekeepers is an important unknown and may be approximately the same number as those registered.

Research aims and objectives

There is very little scientific literature available specific to honey bees in Ireland and only tentative information about their genetic make-up (Jensen *et al.* 2005, Jaffe *et al.* 2010), particularly regarding the native sub-species, *A. m. mellifera*. The research here aimed to take the first steps in elucidating the presence or otherwise of that sub-species on the island of Ireland and towards an understanding of how they have adapted to Ireland's environment following post-glaciation recolonisation (Carreck 2008, Pritchard 2008).

At all times I aimed to consider the bees' ability to withstand infestation by *V. destructor* with either direct effects against the mite or whether adaptations to the local environment, such as in foraging behaviour (Kreitlow and Tarpy 2006), may play a part in colony defence through better colony health (Di Pasquale *et al.* 2013).

The initial research objective was to start a breeding programme and to expand this into the beekeeping community with the aim of determining whether or not, through selective breeding, resistant colony prevalence could be increased in a sample group.

The second part of the research aimed to investigate the existence of colonies that were surviving without assistance from beekeepers. Long term survivors in these free-living colonies were considered likely to provide a reservoir of resistance to varroa.

The final research objective was to build on current light microscopy-based knowledge about the pollen types utilised by Ireland's honey bees through the use of molecular techniques. The aims were to provide a greater understanding of the pollens used and to compare the pollens between and within apiaries. An additional objective was to investigate the bee's

digestive tract microbiome in the same colonies used for pollen research, also using molecular techniques.

The overarching objective linking the studies was to genotype honey bees from each study to gain an understanding of what genotypes were present in Ireland and the extent of introgression from imported conspecifics, facilitating a comparison with the corresponding genotypic picture from other parts of their natural range.

References

Aizen, M. A. and Harder, L. D. (2009) 'The global stock of domesticated honey bees is growing slower than agricultural demand for pollination', *Current Biology,* 19(11), 915-918.

Alburaki, M., Bertrand, B., Legout, H., Moulin, S., Alburaki, A., Sheppard, W. S. and Garnery, L. (2013) 'A fifth major genetic group among honeybees revealed in Syria', *BMC Genetics,* 14, 10pp.-10pp.

Anderson, D. L. and Trueman, J. W. H. (2000) '*Varroa jacobsoni* (Acari : Varroidae) is more than one species', *Experimental and Applied Acarology,* 24(3), 165-189.

Beye, M., Gattermeier, I., Hasselmann, M., Gempe, T., Schioett, M., Baines, J. F., Schlipalius, D., Mougel, F., Emore, C., Rueppell, O., Sirvio, A., Guzman-Novoa, E., Hunt, G., Solignac, M. and Page, R. E. (2006) 'Exceptionally high levels of recombination across the honey bee genome', *Genome Research,* 16(11), 1339-1344.

Boecking, O. and Spivak, M. (1999) 'Behavioral defenses of honey bees against *Varroa jacobsoni* Oud', *Apidologie,* 30(2-3), 141-158.

Breeze, T. D., Bailey, A. P., Balcombe, K. G. and Potts, S. G. (2011) 'Pollination services in the UK: How important are honeybees?', *Agriculture Ecosystems & Environment,* 142(3-4), 137-143.

Carreck, N. L. (2008) 'Are honey bees (*Apis mellifera* L.) native to the British Isles?', *Journal of Apicultural Research,* 47(4), 318-322.

Chauzat, M. P., Cauquil, L., Roy, L., Franco, S., Hendrikx, P. and Ribiere-Chabert, M. (2013) 'Demographics of the European apicultural industry', *Plos One,* 8(11).

Cullen, M. G., Thompson, L. J., Carolan, J. C., Stout, J. C. and Stanley, D. A. (2019) 'Fungicides, herbicides and bees: A systematic review of existing research and methods', *Plos One,* 14(12).

Cuthbertson, A. G. S., Wakefield, M. E., Powell, M. E., Marris, G., Anderson, H., Budge, G. E., Mathers, J. J., Blackburn, L. F. and Brown, M. A. (2013) 'The small hive beetle *Aethina tumida*: A review of its biology and control measures', *Current Zoology,* 59(5), 644-653.

De la Rua, P., Jaffe, R., Dall'Olio, R., Munoz, I. and Serrano, J. (2009) 'Biodiversity, conservation and current threats to European honeybees', *Apidologie,* 40(3), 263-284.

Dejong, D., Roma, D. D. and Goncalves, L. S. (1982) ' A comparative analysis of shaking solutions for the detection of *Varroa jacobsoni* on adult honeybees.1', *Apidologie,* 13(3), 297-306.

Di Pasquale, G., Salignon, M., Le Conte, Y., Belzunces, L. P., Decourtye, A., Kretzschmar, A., Suchail, S., Brunet, J. L. and Alaux, C. (2013) 'Influence of pollen nutrition on honey bee health: Do pollen quality and diversity matter?', *Plos One,* 8(8).

Duffy, J. (2003) 'Ecology and Evolution of Eusociality in Sponge-Dwelling Shrimp', *Evolutionary Ecology of Social and Sexual Systems: Crustaceans as Model Organisms*. DOI:10.1093/acprof:oso/9780195179927.003.0018

Evans, J. D., Aronstein, K., Chen, Y. P., Hetru, C., Imler, J. L., Jiang, H., Kanost, M., Thompson, G. J., Zou, Z. and Hultmark, D. (2006) 'Immune pathways and defence mechanisms in honey bees *Apis mellifera'*, *Insect Molecular Biology,* 15(5), 645-656.

Franck, P., Garnery, L., Solignac, M. and Cornuet, J. M. (2000) 'Molecular confirmation of a fourth lineage in honeybees from the Near East', *Apidologie,* 31(2), 167-180.

Gadau, J., Page Jr, R. E., Werren, J. H. and Schmid-Hempel, P. (2000) 'Genome organization and social evolution in Hymenoptera', *Die Naturwissenschaften,* 87(2), 87-89.

Gill, R. J. and Raine, N. E. (2014) 'Chronic impairment of bumblebee natural foraging behaviour induced by sublethal pesticide exposure', *Functional Ecology,* 28(6), 1459-1471.

Goulson, D. (2013) 'An overview of the environmental risks posed by neonicotinoid insecticides', *Journal of Applied Ecology,* 50(4), 977-987.

Harrington, T., National Apiculture Programme report 2013-16 Department of Agriculture Food and the Marine, Project ref. CPS/Hort/Bees/38

Huang, Z. (2012) 'Varroa mite reproductive biology', *American Bee Journal,* 152, 981-985.

Ilyasov, R. A., Lee, M.-l., Takahashi, J.-i., Kwon, H. W. and Nikolenko, A. G. (2020) 'A revision of subspecies structure of western honey bee *Apis mellifera'*, *Saudi Journal of Biological Sciences,* 27(12), 3615-3621.

Jaffe, R., Dietemann, V., Allsopp, M. H., Costa, C., Crewe, R. M., Dall'olio, R., de la Rua, P., El-Niweiri, M. A. A., Fries, I., Kezic, N., Meusel, M. S., Paxton, R. J., Shaibi, T., Stolle, E. and Moritz, R. F. A. (2010) 'Estimating the density of honeybee colonies across their natural range to fill the gap in pollinator decline censuses', *Conservation Biology,* 24(2), 583-593.

Jensen, A. B., Palmer, K. A., Boomsma, J. J. and Pedersen, B. V. (2005) 'Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honeybee, *Apis mellifera mellifera*, in northwest Europe', *Molecular Ecology,* 14(1), 93-106.

Kreitlow, K. L. and Tarpy, D. R. (2006) 'Environmental and genotypic effects on Russian-hybrid and Italian honey bee (*Apis mellifera*) (Hymenoptera: Apidae) foraging behavior', *Environmental Entomology,* 35(6), 1610-1616.

Meixner, M. D., Costa, C., Kryger, P., Hatjina, F., Bouga, M., Ivanova, E. and Buechler, R. (2010) 'Conserving diversity and vitality for honey bee breeding', *Journal of Apicultural Research,* 49(1), 85-92.

Meixner, M. D., Leta, M. A., Koeniger, N. and Fuchs, S. (2011) 'The honey bees of Ethiopia represent a new subspecies of *Apis mellifera-Apis mellifera simensis* n. ssp', *Apidologie,* 42(3), 425-437.

Nieto, A., Roberts, S.P.M., Kemp, J., Rasmont, P., Kuhlmann, M., García Criado, M.,Biesmeijer, J.C., Bogusch, P., Dathe, H.H., De la Rúa, P., De Meulemeester, T., Dehon,M., Dewulf, A., Ortiz-Sánchez, F.J., Lhomme, P., Pauly, A., Potts, S.G., Praz, C., Quaranta,M., Radchenko, V.G., Scheuchl, E., Smit, J., Straka, J., Terzo, M., Tomozii, B., Window, J.and Michez, D. (2014). European Red List of bees. Luxembourg: *Publication Office of the European Union*.

Nowak, M. A., Tarnita, C. E. and Wilson, E. O. (2010) 'The evolution of eusociality', *Nature,* 466(7310), 1057-1062.

Porto, R. G., de Almeida, R. F., Cruz-Neto, O., Tabarelli, M., Viana, B. F., Peres, C. A. and Lopes, A. V. (2020) 'Pollination ecosystem services: A comprehensive review of economic values, research funding and policy actions', *Food Security,* 12(6), 1425-1442.

Pritchard, D. (2008) Is the dark bee really native to Britain and Ireland, *Bee improvement and conservation* Winter 2008-9 BIBBA

Ramsey, S. D., Ochoa, R., Bauchan, G., Gulbronson, C., Mowery, J. D., Cohen, A., Lim, D., Joklik, J., Cicero, J. M., Ellis, J. D., Hawthorne, D. and vanEngelsdorp, D. (2019) *'Varroa destructor* feeds primarily on honey bee fat body tissue and not hemolymph', *Proceedings of the National Academy of Sciences of the United States of America,* 116(5), 1792-1801.

Requier, F., Rome, Q., Chiron, G., Decante, D., Marion, S., Menard, M., Muller, F., Villemant, C. and Henry, M. (2019) 'Predation of the invasive Asian hornet affects foraging activity and survival probability of honey bees in Western Europe', *Journal of Pest Science,* 92(2), 567-578.

Ruttner, F., (1988) 'Biogeography and taxonomy of honeybees', Springer-Verlag, Berlin; New York.

Sydenham, M. A. K., Eldegard, K. and Totland, Ø. (2014) 'Spatio-temporal variation in species assemblages in field edges: seasonally distinct responses of solitary bees to local habitat characteristics and landscape conditions', *Biodiversity and Conservation,* 23(10), 2393-2414.

Tarpy, D. R. (2003) 'Genetic diversity within honeybee colonies prevents severe infections and promotes colony growth', *Proceedings of the Royal Society B-Biological Sciences,* 270(1510), 99-103.

Tentcheva, D., Gauthier, L., Zappulla, N., Dainat, B., Cousserans, F., Colin, M. E. and Bergoin, M. (2004) 'Prevalence and seasonal variations of six bee viruses in *Apis mellifera* L. and *Varroa destructor* mite populations in France', *Applied and Environmental Microbiology,* 70(12), 7185-7191.

Chapter 1

Investigations into the existence of the Black honey bee (*Apis mellifera mellifera*) in Ireland and the establishment of a breeding programme for *Varroa destructor* resistance

This work formed part of two papers;

Hassett, J., Browne, K. A., McCormack, G. P., Moore, E., Soland, G., Geary, M. and Native Irish Honey Bee, S. (2018) 'A significant pure population of the dark European honey bee (Apis mellifera mellifera) remains in Ireland', *Journal of Apicultural Research,* 57(3), 337-350. and

Henriques, D., Browne, K. A., Barnett, M. W., Parejo, M., Kryger, P., Freeman, T. C., Munoz, I., Garnery, L., Highet, F., Jonhston, J. S., McCormack, G. P. and Pinto, M. A. (2018) 'High sample throughput genotyping for estimating C-lineage introgression in the dark honey bee: an accurate and cost-effective SNP-based tool', *Scientific Reports,* 8.

Also of two short articles “The search for tolerance to Varroa” and “The native Irish honey bee breeding project. The first year!” both in the *Four Seasons magazine* periodical magazine of the Native Irish Honey Bee Society.

(For Hassett *et al* 2018, my involvement included sample collection, DNA preparation and PCR of mtDNA, data analysis, assisting in production of phylogenetic networks and writing.
For Henriques *et al* 2018, my involvement included sample collection, DNA preparation, raw data reformatting, data analysis, writing the introduction and editing.)

INTRODUCTION

Acaricide treatments to control the obligate honey bee parasite *Varroa destructor* (Anderson and Trueman 2000) (A.K.A. "varroa") have been developed and used for over 50 years. Despite this varroa persists as a primary source of colony losses worldwide with some chemical treatments experiencing reduced efficacy due to adaptation by the mite (Martin 2015) or difficulties achieving the precise ambient conditions to allow them to act at their optimum (Rosenkranz *et al*. 2010). Furthermore, the potential human health risks posed during the application of these poisons and their accumulation in edible hive products (Smodiš Škerl *et al.* 2010) are troubling for beekeeper and consumer. Consequently, beekeepers are beginning to see greater utility in the natural social immunity characteristics of colonies to help ameliorate losses (Blacquière and Panziera 2018). In Ireland and elsewhere, beekeepers are actively seeking tolerant colonies within their own populations, and some are claiming long-term survival of untreated colonies (McMullan 2018, Pritchard, 2018, Boerjan *et al*. 2018). Whilst the validity of many claims of resistance to varroa, particularly the purely anecdotal evidence presented by beekeepers, have not been independently verified, the evidence is compelling and worthy of further investigation.

Expertise in selective breeding exists within the beekeeping community of both Northern Ireland and the Republic of Ireland (hereafter referred to as Ireland) for colony traits sought by both commercial and hobbyist beekeepers alike. Within the membership of the Native Irish Honey Bee Society (NIHBS), which has members across the entire island of Ireland, hobbyist-level selection for genetically pure *Apis mellifera mellifera* exists alongside more rigorous breeding programmes that include artificial insemination (AI).

Breeding programmes for tolerance/resistance to *V. destructor,* from which to adopt techniques, exist elsewhere but not yet in Ireland. The Abeitsgemeinschaft Toleranzzucht (AGT) in Germany and the related international Smartbees programme test specifically for hygienic behaviour which is known to reduce the mite reproduction rate (Ibrahim and Spivak 2006), using either a pin-kill (PKB) or freeze-kill brood (FKB) assay. Data from the assay is combined with other colony characteristics and takes due consideration of environmental effects, to provide a breeding value for each colony. From these data a breeding value is determined and colonies with the highest breeding values are used to produce queens and drones for mating stations. In Russia's Primorsky region, non-native *A.m.mellifera* were imported for honey production which resulted in them living in close proximity to *A. cerana,*

the original host of *V. destructor*. As a result of an extended period of exposure to the mite, they have developed a genetic based resistance thought to comprise varroa specific hygienic behaviour (VSH) and the less understood suppression of mite reproduction (SMR) (Rinderer *et al.* 1997). Attempts have been made to introduce the resistance characteristics into other populations. The United States Department of Agriculture (USDA) have successfully used the Russian Primorsky honey bees to develop colonies that exhibit resistance to varroa through both characteristics (Harris *et al.* 2002, Rinderer *et al.* 2010) and similar use has been made in Europe (Büchler *et al.* 2010). However, underlining a general lack of understanding about how characteristics are genetically cross-linked within the honey bee genome, both the US and German programmes experienced unwanted colony characteristics such as unsatisfactory brood pattern, increased defensive behaviour and lowered productivity (Rinderer *et al.* 2010, Büchler *et al.* 2010) displaying a possible weakness in the targeting of specific resistance characteristics only.

The local apiculture conditions under which these programmes operate differ considerably from the situation in Ireland, meaning that knowledge transfer is also not straightforward. Germany has all but replaced its native *A. m. mellifera* with *A. m. carnica* and in the Americas the honey bee is not native. In contrast, Ireland's beekeepers have been adamant that pure *A. m. mellifera* persists in the country, a claim which, if proven, could make Ireland one of the last holdouts for the sub-species within its natural range, given the consequences of varroa infestation and bee importation in mainland Europe (Garnery *et al.* 1998a, Garnery *et al.* 1998b, De la Rua *et al.* 2009). Compared to Ireland there is a greater emphasis on commercial beekeeping in Germany and the USA which means that the AGT and USDA breeding programmes need mainly concern themselves with creating a resistant strain, albeit with agreeable handling and production characteristics, and sufficient genetic diversity. Sub-species conservation is not at the core of these programmes.

At the start of this work NIHBS had begun to develop the idea of a breeding programme for varroa resistance/tolerance using the sugar shaker method of selection (Scott, 2014). This method of varroa infestation assessment is cheap, simple to perform and accurate (Macedo *et al.* 2002, Lee *et al.* 2010). Furthermore, it is non-toxic and non-lethal to bees which would hopefully enhance its uptake rate amongst beekeepers. This study aimed to instigate and to provide scientific support to the above breeding programme with its success or otherwise serving to inform future efforts. To some extent the programme aimed to mimic the AGT and

Smartbees programmes through collection of data on desirable colony traits but using criteria developed here with some Irish beekeepers.

An essential part of any breeding programme where the origin of the bees is of high importance, as *A. m. mellifera* is to NIHBS, is the genotyping of participating colonies. This work aimed to discover the sub-species present and the extent of hybridization within the honey bee population managed by NIHBS members and to build on the limited evidence of *Apis mellifera mellifera's* existence in Ireland (Jensen *et al.* 2005). The original premise was for the test programme to be limited to approximately 30 experienced beekeepers who would provide bees from their hives, test them for varroa load and leave them untreated as indicated. Queens produced from putative resistant colonies would be used to create a line of resistant colonies in a research apiary at NUI Galway and a second, susceptible, lineage would be created for comparison. The genotyping was initially intended to start as a way of searching for genotypic differences between the resistant and susceptible lineages however the subsequent lack of low-varroa untreated colonies and inconsistency in data collection drove the research a different direction.

MATERIALS AND METHODS

Selection of participants

Participants in 2013 and 2014 were mostly NIHBS committee members with the intention to field test the data collection format and provide preliminary data.

In 2015 participants were sourced using direct email and phone contact from a list supplied by Dr Michael Geary (Limerick Institute of Technology, Limerick, Ireland) of beekeepers who had expressed an interest to him in a breeding programme. These came from within and outside the NIHBS membership. A maximum of thirty participants were sought, consisting of beekeepers who had individual apiaries containing ten or more colonies. Details of the methodology were sent to all those that expressed an interest and those who proceeded further were considered to have self-determined both their availability and experience to commit to the aims and methods of the programme.

For the 2016 data collection, NIHBS contacted their entire membership directly by email and included all necessary details to allow participation by any willing volunteers. A similar beekeeper self-assessment operated as per 2015.

Assessment of mite load and breeding colony selection

Participants were instructed to randomly select eight to ten colonies from a single apiary to provide data representative of the apiary's mean mite load (Lee *et al.* 2010). Where less than eight colonies were present, the entire apiary was to be assessed. Participants were asked to perform an assessment of the mite load in each colony in May and August of each year although these times could vary to ensure the assessment took place before any form of chemical or biotechnical treatment for varroa. Mite load was assessed using the "sugar shake test" using icing sugar (Appendix BP I, Pg.42) added to a sample of approximately 300 bees (100mL) to encourage auto-grooming and dislodgement of the mites (Macedo *et al.* 2002). The mites were counted and converted to a load percentage. The presence of sealed brood on the test frame(s) that the bees were shook from required the mite count to be doubled (Lee *et al*. 2010). In addition, the proportion of sealed and open brood present on the test frame, the age and provenance of the queen (to facilitate tracing to other resistant colonies) and previous treatment for varroa were recorded on a provided form (Appendix BP II, Pg.45). Colonies with varroa load of ≤2% were to be left untreated and used preferentially for breeding queens. Colonies above that threshold were permitted treatment per the

beekeepers' standard management methods and were to be re-queened at a suitable time with daughters of below-threshold queens obtained from within the same apiary or from other participating beekeepers.

Evaluation of traditional traits

Colony characteristics based on the ordinal scale developed by the Galtee bee breeding group (GBBG) were recorded (Appendix BP II) to help inform decisions on selection and avoid the potential loss of desirable traits. Docility, steadiness on the comb, brood pattern, pollen storage and comb building were evaluated and scored by the beekeeper against descriptive ratings from "unsatisfactory" to "outstanding" (0 to 5 respectively). A similar evaluation scale for propolis coverage was added in light of its positive association with hygienic colonies (Nicodemo *et al.* 2013).

Transmission of the project methodology to participants

At the beginning of each year, participants were given introductory information about the methodology, aims and objectives of the programme (Appendices BP IV and BP V, Pgs. 49 & 51). Acting on feedback from earlier participants the introductory information was eventually accompanied by an FAQ sheet (Appendix BP VI, Pg. 52) to help reduce errors in applying the methodology and time dealing with administration. Participants were also given a data recording form with explanatory notes, which had undergone various iterations to reduce complexity following discussion with beekeepers (Appendices BP II and BP III, Pg.47). An explanation of the sugar shaker assay in written/pictorial form was made available (Appendix BP I). Regional representatives for NIHBS conducted practical demonstrations of the method in beekeeping association (BKA) apiaries, researcher talks and articles in beekeeping periodicals and talks about the programme were undertaken along with demonstrations of the methodology. To promote expansion of the breeding programme within the beekeeping community all of the forms were freely available for downloading from the NIHBS website (www.nihbs.org) which also provided a video of the sugar shaker assay.

Data handling and analysis

Incoming data for each year were recorded in Excel format in a single file. Summary data were extracted manually, and summary statistics were calculated using the dedicated formulae within Excel.

Genotyping

DNA extraction

Participants submitted worker bees from each colony by post which were stored at -20°C until required. DNA was extracted from the hind legs of two bees per colony using the standard protocol of the E.Z.N.A. DNA extraction kit (Omega bio-tek, 2013). The extracted DNA was kept individually. Data from additional beekeepers were processed by Jack Hassett (University of Limerick) in a related project.

Mitochondrial Data

The mitochondrial COI-COII intergenic region of 129 bees, representing 75 colonies from 38 beekeepers, was amplified by PCR using Illustra PuReTaq Ready-to-go PCR beads in a 25 µL reaction consisting of 10 pmol each (2.5 µL at our concentration) of primers E2 and H2 (Garnery *et al.* 1992), 10 µL of template DNA and 10 µL of bidest (double distilled) water. Reactions consisted of an initial denaturation of 5 minutes at 95°C followed by 35 cycles of: 94 °C for 45 secs, 48 °C for 45 secs and 62 °C for 2 mins, with a final extension of 20 mins at 65 °C.

Successful amplification was determined using electrophoresis via a 1% Agarose (Sigma) gel and the PCR products were purified using GeneJET PCR purification kit (Thermo Scientific).

Sequences were generated by LGC Genomics, Germany, using Sangar sequencing. The chromatogram for each sequence was assessed using MEGA6 (Tamura *et al.* 2013) and subjectively assigned a rating based on trace clarity and signal strength (Appendix BP VII, Pg. 53) to inform decisions on ambiguities in base calling. When forward and reverse sequences, where required, had been assessed, a consensus sequence was imported into a multiple alignment with other Irish sequences forming part of the wider study by Hassett *et al.* (2018) and European representative sequences downloaded from GenBank. The intergenic sequence consists of a P sequence (54bp, 100% A & T) and repetitions of a Q sequence (196bp, 93.4% A & T) (Cornuet *et al.* 1991). The similarity and repetitiveness of the Q sequences

imposed limitations on accuracy of the automated alignment protocols (ClustalW and MUSCLE) in MEGA6, therefore all sequences were aligned by eye. A reference alignment containing representative sequences from the most likely subspecies present in Ireland was created using sequences available in GenBank (sequences from *A. m. ligustica, A. m. carnica, A. m. iberiensis, A. m. scutellata*). This alignment was used to screen sequences from Irish bees to determine sub- species status of their mtDNA. Subsequently all those identified as *A. m. mellifera* were aligned to all available European *A. m. mellifera* sequences which were those from Rortais *et al*. (2011) and Pinto *et al*. (2014). Unfortunately, sequences from Jensen et al. (2005) were not available. Identical sequences were identified as were new variants of the region from Irish bees. Only one representative sequence for each mitotype was retained in the alignment. Phylogenetic networks were constructed under statistical parsimony using TCS 1.21 (Clement *et al*. 2000).

Microsatellite Data and Single Nucleotide Polymorphisms (SNPs)

The DNA from 172 bees were sent to EcoGenics (Switzerland) to be genotyped using 12 informative microsatellite markers (A007, A28, A29, A43, A76, A273, Ac306, Ap1, Ap33, Ap226, Ap289, B24) (Estoup *et al*. 1995, Garnery *et al*. 1998, Soland-Reckeweg *et al*. 2009, Alburaki *et al*. 2013, Meixner *et al*. 2013). Bayesian analysis and visualisation of population assignment between C and M lineages was conducted in Structure V2.3.4 (Pritchard *et al.* 2000) by Jack Hassett (Limerick Institute of Technology) and included in a larger study (Hassett *et al*. 2018).

DNA from 46 bees were genotyped using a 127 SNP assay split over four panels, at Instituto Gulbenkian Ciência (Portugal) and using a combination of Agena BioScience's iPLEX and Sequenom's MassARRAY™ MALDI-TOF. The resultant data were analysed in collaboration with Instituto Politécnico de Bragança (IPB) (Portugal) and formed part of Henriques *et al*. (2018). C-lineage to M-lineage proportion assignment (Q-values) for K=2 was conducted using ADMIXTURE V1.23 (Alexander *et al*. 2009). CLUMPAK (Kopelman *et al*. 2015) was used to summarise and visualise the Q values.

RESULTS

Varroa mite load

Sugar shaker tests for mite load were conducted once each in 2013, 2014, 2015 and twice (Spring and Autumn) in 2016. Between Autumn 2013 and Autumn 2016, 51 different beekeepers provided data returns for 419 different colonies from 97 apiaries (Table 1).

Table 1: Mite load data, expressed as median/mean mite percentages, for the five testing periods from Autumn 2013 to Autumn 2016 at three different "levels" (All colonies, all apiaries and only apiaries with five or more colonies) SD= Standard deviation; Beeks=beekeepers.

a) Colony level

Testing period	No. of beekeepers	No. of colonies	Median mite %	Mean mite %	SD
Aug/Sept 2013	2	21	1.00	1.79	3.01
Sept/Oct 2014	9	98	2.00	9.94	10.98
June-Aug 2015	14	116	2.00	4.55	6.98
Spring 2016	37	226	0.67	1.71	3.22
Autumn 2016	26	172	3.33	5.35	6.93

b) Apiary level 1 (all)

	No. of apiaries	Mean colonies/apiary	SD	Mean mite %	SD
Aug/Sept 2013	3	7.00	4.00	1.38	1.13
Sept/Oct 2014	23	4.41	3.38	9.87	13.5
June-Aug 2015	24	4.79	3.27	5.30	4.56
Spring 2016	54	4.26	2.44	1.86	2.15
Autumn 2016	39	4.41	2.74	5.62	5.15

c) Apiary level 2 (≥5 colonies)

	No. of apiaries	Mean colonies/apiary	SD	Mean mite %	SD
Aug/Sept 2013	2	9.00	2.83	2.02	0.36
Sept/Oct 2014	8	7.60	2.83	6.60	7.53
June-Aug 2015	12	7.91	1.81	4.11	3.54
Spring 2016	20	6.90	1.52	1.43	1.42
Autumn 2016	17	7.00	1.94	5.08	3.87

Testing period 2013 only had three beekeepers involved as a test of the data acquisition method. For all colonies (Table 1, a) the mite percentages ranges were from 0.00% (every test) to 14.00%, 52.00%, 42.00%, 24.00% and 48.00% respectively from 2013 to 2016. The high standard deviations for the mean mite percentage loads reflected the wide value ranges. Median mite percentages were calculated for the same tests to reduce the skew in the data

caused by a small number of extremely high mite counts and showed considerably lower values.

The mean number of colonies per apiary (Table 1, b) differed slightly for the 2014 to 2016 periods with a range of 1 to 13 colonies/apiary. When only the data from apiaries with five or more colonies (35% to 50% of the complete dataset per year) were analysed (Table 1, c) the mean colonies/apiary increased by approximately 3 colonies.

Threshold mite load percentages

Across the five test periods, between 45% and 81% of the colonies had mite loads that were less than or equal to the 2.00% threshold mite load (Table 2) required to enable beekeepers to avoid chemical acaricide treatments.

Table 2: The proportion, by percentage, of all colonies for each test period which returned mite loads of zero, ≤2% or ≥10%. Beekeepers with colonies having values of ≤2% were encouraged to not chemically treat these colonies as part of the breeding programme. Colonies with a mite load of <10% are generally still considered as having a low varroa load.

Varroa load percentages for all colonies			
Testing period	**0% load**	**≤2% load (incl. 0%)**	**≥10% load**
Aug/Sept 2013	19.50%	80.95%	4.76%
Sept/Oct 2014	15.31%	52.04%	23.47%
June-Aug 2015	17.24%	50.86%	13.79%
Spring 2016	43.81%	76.99%	3.09%
Autumn 2016	18.02%	44.77%	16.86%

Of the more widely sampled years (2014 to 2016) 2014, 2015 and Autumn 2016 showed similar proportions of colonies with mite loads ≤2% (45% to 52%) whereas Spring 2016 differed considerably with 77% under the treatment threshold.

Data tracking

Varroa loads for individual colonies were tracked across the five testing points from 2013 to Autumn 2016 (Appendix BP VIII) and summarised (Table 3). From the initial 21 colonies tested in 2013, 81% had ≤2% load and although 15 (71%) were also tested in 2014, only colony B6W-H1 (Table 4) had continuous data through to 2016b. From 2013 to 2015, 38% of the colonies had continual data. Forty percent of colonies tested in 2014 were also tested in 2015. However, 13% had continued testing to 2016a and <1% to 2016b. From the 116 colonies tested in 2015, 13% provided data for 2016a and 3% were continuous to 2016b.

No untreated colonies had more than one test except colonies DR4 and DR5 from beekeeper B7T which had data trackable from 2015 to 2016b (Table 4). These two colonies recorded mite percentage loads across the three tests (2015, 2016a & 2016b) as follows DR4: 4.00, 1.33 & 2.00, DR5: 3.00, 0.67 & 4.00. For each of these two colonies the middle one of the three tests was recorded as occurring after a period ≥12 months without treatment.

For the test periods which had significant sample sizes (2014, 2015, 2016a and 2016b), 39 colonies from 2014 were also tested in 2015, reducing to 13 of those colonies in 2016a and one colony in 2016b. The 39 colonies entailed 40% of the 2014 total but only 34% of the total tested in 2015, 6% of 2016a and 1% of 2016b. Although 18 (46%) of the 39 tested colonies with mite loads ≤2% were tested in both 2014 and 2015, this reduced to 7 colonies in 2016a and 1 colony in 2016b. Additionally none of these low mite load colonies were untreated.

Four beekeepers from four counties provided mite load data on 25 colonies from six apiaries for three or more periods (Table 4). However, these data periods were not necessarily continuous. Although the test periods for which data were supplied do not overlap across all colonies, 19 colonies had three continuous test results. Of these, one had results from four continuous tests and one from five tests. All of these colonies had been treated for varroa in the 12 months prior to each.

The mite loads ranged from 0.00% to 52.00% across the group of 25 colonies although the higher values were obtained from only five colonies. Four of those five colonies (141, DA8, G4 and PD1) were from two apiaries located in urban settings and the location of the fifth, B2, was very rural. The mite load of colony B2 reduced from 14.67% in June-August 2015 to 1.33% in Spring 2016. In 2016 colonies 141 and G4 in different Antrim apiaries belonging to B2Am showed mite loads that had reduced to sustainable levels of 4.00% and 0.00% from 2014 highs of 46.67% and 52.00% respectively. Colony DA8 showed a similar reduction on the same timescale but a more gradual improvement. The median load percentages for this cohort across all test periods ranged from 0.33% to 2.67%.

Table 3: How the numbers of tested colonies carried on "from" one test date "to" the next and subsequent tests. The table shows the number of colonies with continual varroa load data. Section A) tracks now the colonies that had data recorded in 2013 were carried through into 2016b (Autumn 2016). Section B) is from 2014 to 2016b, C) 2015 to 2016b and D) 2016a to 2016b. The proportion of the test that the colonies were assessed (the "from" test) and their proportion of the total colonies for the "to" test are also shown. The numbers and proportions of colonies which were untreated in the 12 months prior to the varroa assay and with varroa loads ≤2% are also shown. Total colonies per test were 21 in 2013, 98 in 2014, 116 in 2015, 226 in 2016a and 172 in 2016b.

	Section A				
Tests "from" "to"	2013	2013 to 2014	2013 to 2015	2013 to 2016a	2013 to 2016b
Number	**21**	**15**	**8**	**3**	**1**
Proportion of "from" carried	-	0.71	0.38	0.14	0.05
Proportion of 2013 in "to" test	-	0.15	0.07	0.01	0.01
Number untreated	**0**	**0**	**0**	**0**	**0**
Number ≤2%	**17**	**3**	**4**	**3**	**1**
Proportion of "from" ≤2%	0.81	0.2	0.5	1.00	0.05
Proportion carried of "to" ≤2%	-	0.03	0.03	0.01	0.01
			Section B		
Tests "from" "to"			2014 to 2015	2014 to 2016a	2014 to 2016b
Number			**39**	**13**	**1**
Proportion of "from" carried			0.40	0.13	0.01
Proportion of 2014 in "to" test			0.34	0.06	0.01
Number untreated			**0**	**0**	**0**
Number ≤2%			**18**	**7**	**1**
Proportion of "from" ≤2%			0.46	0.03	1.00
Proportion carried of "to" ≤2%			0.16	0.03	0.01
				Section C	
Tests "from" "to"				2015 to 2016a	2015 to 2016b
Number				**16**	**3**
Proportion of "from" carried				0.13	0.03
Proportion of 2015 in "to" test				0.07	0.02
Number untreated				**0**	**2**
Number ≤2%				**9**	**2**
Proportion of "from" ≤2%				0.56	0.67
Proportion carried of "to" ≤2%				0.04	0.01
					Section D
Tests "from" "to"					2016a to 2016b
Number					**113**
Proportion of "from" carried					0.50
Proportion of 2016a to total of "to"					0.66
Number untreated					**11**
Number ≤2%					**36**
Proportion of "from" ≤2%					0.32
Proportion carried of "to" ≤2%					0.19

Table 4: Mite load percentages for 25 colonies from six apiaries for which data was provided for three or more of the five test periods. Four beekeepers, one each from counties Galway, Antrim, Waterford and Tipperary were represented. 2016a and 2016b are Spring and Autumn 2016 respectively. 14 colonies had continuous data for three tests, two for four tests, one continuous for five tests. Four were discontinuous for three tests and four discontinuous for four tests.

Beekeeper	Colony	2013	2014	2015	2016a	2016b
B1G	B2		4.00	14.67	1.33	
"	B3		2.67	2.00	2.67	
"	B5		1.33	0.00	3.33	
"	B6		0.00	2.00	0.00	
"	K1		2.00	4.00	4.00	
"	K2		2.67	3.33	4.67	
"	K4		6.67	2.67	0.67	
"	K5		2.00	1.33	2.67	
B2Am	94	0.00	12.00	4.00	0.67	
"	141		46.67	3.33	4.00	
"	BC1	3.00	2.67		2.67	
"	DA8	0.00	20.00	17.33	0.00	
"	G4		52.00	2.67	0.00	
"	PD1	0.30	26.67	20.67		
B6W	H1	1.00	0.00	1.33	0.67	1.33
"	H2	2.00	4.00	1.33		0.00
"	H3	0.00	0.67		1.33	0.00
"	H4	0.00	2.67	4.00		0.00
"	H5	1.00	2.67			1.33
"	H6	1.00	0.67	1.33		0.67
"	H9	1.00	5.33	0.00		
B7T	DA1		2.00	1.53		0.00
"	DA2		3.33	2.87		0.00
"	DR4			2.67	1.33	2.00
"	DR5			2.00	0.67	4.00
Median		**1.00**	**2.67**	**2.67**	**1.33**	**0.33**

Acaricide treatment analysis

In the four test periods from 2014 to 2016 thymol and oxalic acid-based compounds constituted between 68.15% and 90.77% of the acaricidal treatments used in the twelve months prior to testing. They were used either singly, in tandem or at different times of year, usually thymol in Autumn and oxalic acid in spring. In 2013 100% of the colonies were treated using oxalic acid.

Apiguard™ or ApilifeVar™ were used for thymol treatments and Apiboxal™ or oxalic acid crystals used for oxalic acid treatments. Commercially available forms of formic acid and flumethrin and fluvalinate were used to a lesser extent.

In 2014 approximately 26.53% of colonies did not have their treatments recorded, 68.10% went unrecorded in 2015 and 4.42% and 21.51% respectively for the spring and autumn 2016 recordings.
The percentage of colony recorded as untreated in the previous 12 months were 0% (2013), 11.11% (2014), 18.92% (2015), 1.39% (2016a) and 14.07% (2016b).

Overall, the mite percentage data showed considerable variation between colonies within a given test period and also for individual colonies between periods. Although there were un-treated colonies recorded, particularly in 2016a, with low or zero varroa load, there was a lack of continuous data on such colonies.

Genotyping

Bees supplied by 47 beekeepers from 22 counties (Antrim, Armagh, Carlow, Cork, Derry, Dublin, Galway, Kerry, Laois, Leitrim, Limerick, Louth, Mayo, Monaghan, Offaly, Roscommon, Sligo, Tipperary, Tyrone, Waterford, Wexford and Wicklow) were genotyped using one or more of MtDNA, microsatellite and SNP data (Table 5). Thirty-two of these beekeepers were participants in the breeding programme and fifteen had only supplied bees to facilitate a geographic spread of the genotyping data.

Table 5: Distribution of molecular data from 32 breeding programme beekeepers and 15 others added to provide best geographic coverage. Both the mitochondrial (mtDNA) and microsatellite (M'sats) data formed part of Hassett *et al.* (2018). The SNPs data formed part of Henriques *et al.* (2018).

Microsatellite Bees column = Number used (Number also used in mtDNA are in brackets)

SNPs Mean Q values: *= Genotypes that were below the threshold for assignment to M lineage.

SNPs Bees: # = Same bees as used in microsatellite genotyping

Figures in brackets below totals are the number of beekeepers represented by each genotyping method.

Mitochondrial group(s): [1] - Differing genotypes produced from one colony. See note in text

		Mitochondrial			Microsatellite			SNPs		
B'keeper	**County**	**Bees**	**Colonies**	**Group(s)**	**Bees**	**Colonies**	**Mean Q value**	**Bees**	**Colonies**	**Mean Q value**
B1G	Galway	4	2	PQQ/PQQQ [1]	1	1	0.998			
B2Am	Antrim	7	4	PQQ/PQQQQ	14 (6)	9	0.995	3 #	3	0.970
B3WW	Wicklow	2	1	PQQ	2	1	0.998	1 #	1	0.987
B4C	Cork	6	4	PQQ	7	4	0.995	3 #	3	0.989
B5W	Waterford	2	1	PQQ						
B6W	Waterford	11	7	PQQ	9	7	0.996	3 #	3	0.996
B7T	Tipperary	5	5	PQQ	10	5	0.995	4	4	0.995
B8G	Galway	3	2	PQQ				1	1	0.988
B9Ah	Armagh	3	2	PQQ	2	1	0.994	1 #	1	0.867*
B10MO	Mayo	3	2	PQQ						
B11C	Cork	2	1	PQQ	4	2	0.998			
B12WX	Wexford	6	4	PQQ	4	3	0.994	2 #	2	0.988
B13C	Cork				1	1	0.945			
B15G	Galway	2	1	PQQ				1	1	1.000
B16W	Waterford	1	1	PQQ	2	1	0.999			
B17G	Galway	2	1	PQQ						
B18WX	Wexford	1	1	PQQ	2	1	0.994			
B19G	Galway	3	2	PQQ				2	2	0.989
B20MO	Mayo				3	1	0.998	1	1	0.991
B21MO	Mayo				1	1	0.997			
B22D	Dublin				2	1	0.999			
B23G	Galway				2	1	0.997			
B24RN	Roscommon				1	1	0.997	1 #	1	0.988
B25LM	Leitrim				2	1	0.974			
B26CW	Carlow	7	4	PQQ	9 (2)	5	0.998	2 #	2	0.988
B27OY	Offaly	2	1	PQQ	2 (2)	1	0.999	1 #	1	1.000
B28KY	Kerry	3	2	PQQ	2 (2)	1	0.993	1 #	1	0.990
B30D	Dublin	2	1	PQQ	2 (2)	1	0.990	1 #	1	0.989
B31LM	Leitrim	2	1	PQQ	2 (2)	1	0.992	1	1	0.991
B32OY	Offaly	2	1	PQQ	2 (2)	1	0.990	1 #	1	0.970
B33LH	Louth	4	3	PQQ	20	11	0.997	2	2	0.984
B34T	Tipperary	2	1	PQQQ	2 (2)	1	0.996			
B35KY	Kerry							1	1	0.997
B37OY	Offaly	2	1	PQQ	2 (2)	1	0.997			
B38Ah	Armagh	2	1	PQQ	2 (2)	1	0.991			
B41WX	Wexford	6	3	PQQ	10	6	0.997	2 #	2	0.996
B42G	Galway	2	1	PQQQ	2 (2)	1	0.967	1 #	1	0.991
B44LS	Laois	4	2	PQQ	4 (4)	2	0.999	1 #	1	1.000
B54D	Dublin	2	1	PQQ	2 (2)	1	0.993			
B55Dy	Derry							2	1	0.981
B56SO	Sligo	1	1	PQQQQ	2 (2)	1	0.998	1 #	1	0.981
B58T	Tipperary	6	5	PQQ	22	13	0.995			
B61W	Waterford	9	6	PQQ	12	7	0.993	3 #	3	0.997
B63L	Limerick	2	1	PQQQ	2 (2)	1	0.994	1 #	1	1.000
B67KY	Kerry	2	1	PQQQ	1 (1)	1	0.994			
B68MN	Monaghan	2	1	PQQ	2 (2)	1	0.988	1 #	1	0.867*
B69Te	Tyrone	2	1	PQQ	2 (2)	1	0.961	1 #	1	0.939
47	**22 unique counties**	**129** *(38)*	**80**		**175** *(39)*	**101**		**46** *(29)*	**45**	

Mitochondrial results

For 80 colonies from 38 beekeepers (Table 5) 129 COI – COII intergenic sequences were successfully produced. All included a P element (Garnery *et al.* 1992) indicating M-lineage queens for all of the colonies tested. 73 colonies were PQQ mitotypes although one colony (B1G in Galway) (Table 5, Note[1]) produced different mitotypes from the two bees, PQQ and PQQQ respectively, possibly as a consequence of the chance collection of a robber bee. Four colonies from apiaries in Galway, Limerick, Kerry and Tipperary were PQQQ mitotype and a further two colonies from Sligo and Antrim were PQQQQ.

The breeding programme (BP) groups' mitotypes were analysed to variant level and analysed with other European mitotypes (Figure 1) as part of the wider study by Hassett *et al*. (2018).

PQ mitotype

No PQ mitotype was discovered.

PQQ mitotype

Sixty percent of all Irish bees sequenced were identical to the PQQ M4d and M4e variants from the Netherlands namely, KF274627M4d KF274628M4e from Pinto *et al*. (2014). There were eighteen different Irish PQQ mitotype variants and the BP group were represented in eight of these. Three BP group beekeepers, B44LS, B37OY and B69Te, had colonies with unique variants, differing by one or two mutational steps from the M4e Netherlands variant which was used as the baseline sequence.

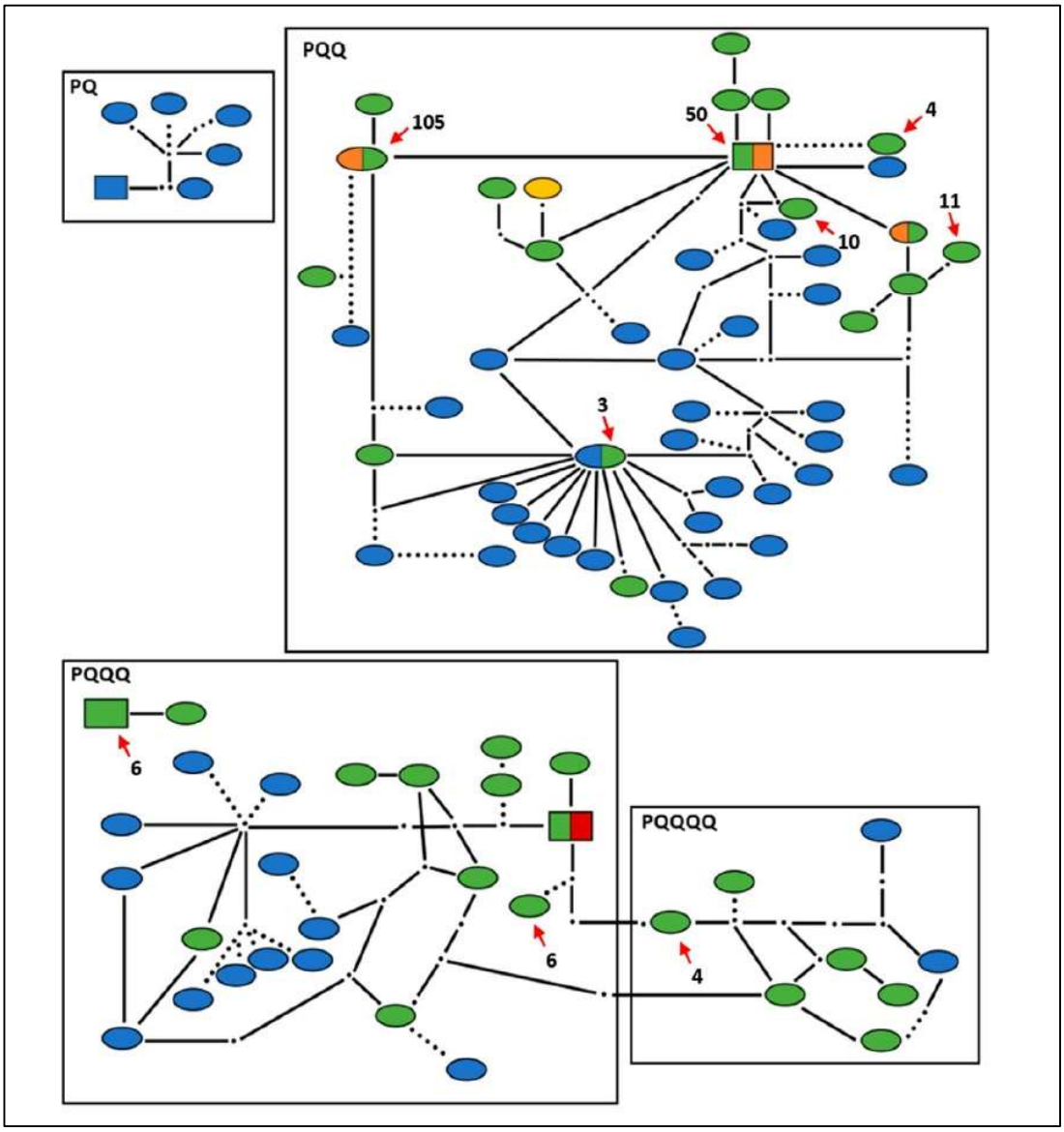
PQQQ mitotype

The BP group had representative colonies in five of the 12 Irish PQQQ mitootype variants identified. Three of the five, B34T, B1G and B42G, had unique variants.

PQQQQ mitotype

Two variants which were unique to the BP group, representing three different colonies (two colonies from beekeeper B2Am and one from B56S), were identified from the six Irish variants identified.

Figure 1: Mitotype variant network for the MtDNA COI-COII intergenic region of Irish (green), Dutch (orange), UK (red) and other European (blue) *Apis mellifera*. An *A.m.ligustica* type (yellow) included from Pinto *et al.* (2014). The 129 sequences produced from the breeding programme colonies are included within the Irish total of 255 but are not delineated. Boxes representing multiple Irish samples of identical mitotype are indicated by numbers and directed by red arrows. (Adapted from Hassett *et al.* 2018)



Microsatellite results

From 101 colonies, 175 bees were successfully genotyped (Table 5). This represented bees from 39 beekeepers. There were seven new beekeepers and six lost, relative to those that participated in the mitochondrial data, giving an 84% participant overlap.

For 46.15% (n=18) of beekeepers the bees used for both microsatellite and mitochondrial genotyping were identical (in the case of beekeepers B2Am and B26CW additional bees were also included in the microsatellite data)

The data were combined with reference populations (*A. m. mellifera* Sweden (n = 6), *A. m. mellifera* France (n = 24), *A. m. mellifera* Norway (n = 18), *A. m. mellifera* Switzerland (n = 17), *A.m. ligustica* Italy (n = 55), *A. m. carnica* Austria (n = 62), *A. m. carnica* Slovenia (n = 21), *A. m. carnica* Switzerland (n = 91)) and analysed in Structure (Figure 2). The analysis indicated a lack of introgression in the bees from the breeding programme colonies with all being above the 0.90 minimum Q-value for assignment to the M-lineage (Table 5).

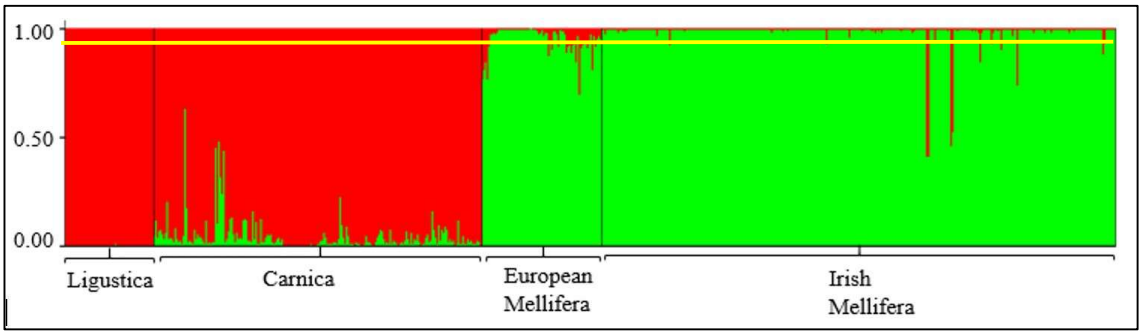


Figure 2: STRUCTURE (K=2) analysis of 12 microsatellite markers for Irish and European honey bees. Individual bees are represented by a single vertical line. Vertical black lines delineate the input putative populations. The 175 breeding programme bees are included with other Irish bees within the Irish Mellifera section. Structure analysis produces assignment values to both the M (green) and C (red) lineages from 0.0 to 1.0. as represented here by individual lines. An assignment value of ≥0.90, indicated by the horizontal yellow line, indicates purity to that lineage. Seven Irish bees, none of which came from the breeding programme, have a red portion of their vertical line that extends below the 0.90 value yellow line, indicating they have C-lineage introgression.

SNPs results

Forty-six bees from 29 beekeepers, representing 45 colonies were successfully genotyped using SNPs (Table 5). From the Admixture analysis, 44 (95.65%) of the bees were assigned as M-lineage, with values above the 0.90 minimum assignment Q value (Table 5). Thirty-one of the DNA extractions used were identical to those used for microsatellite genotyping and of those 31, 13 had also been used for mitochondrial genotyping.

Comparing the lineage assignments between microsatellite (Structure) and SNPs (Admixture), 29 of the 31 bees were assigned as M-lineage by both methods. Two bees, B9Ah145 and B68Mn213 were both assigned as M-lineage by microsatellite analysis and C-lineage using SNPs.

The data from the 46 bees formed part of an analysis of the introgression levels between M and C lineages in European countries (Figure 3). Ireland showed very low lineage introgression relative to the other nine countries assessed. None of the Irish samples were obtained from areas where honey bee mating is controlled or protected.

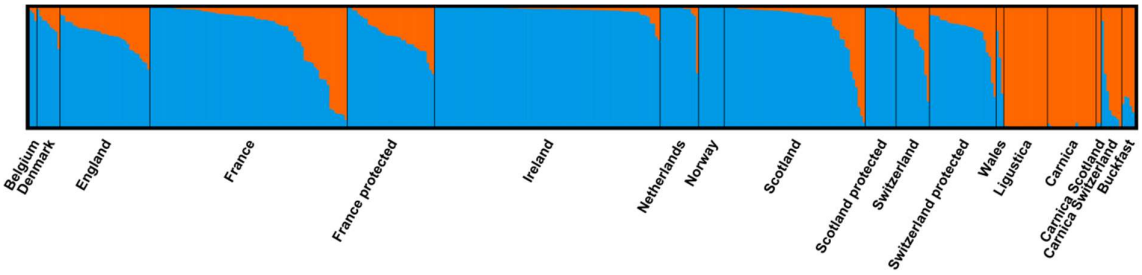


Figure 3. ADMIXTURE (K=2) analysis of 117 SNPs using individuals from 10 different countries where M-lineage (blue) *A. m. mellifera* is native including Ireland where the 46 breeding programme bees are represented in a total of 85 Irish bees. Italy, Croatia and Serbia represent the native range of the C-lineage (orange) *A. m.carnica* and *A. m. ligustica*). Each bar represents one individual with Q values assigned to M and C lineage ranging from 0.0 to 1.0. Vertical black lines delineate the input putative populations. An assignment of ≥0.9 to either lineage indicates purity to that lineage. 91% of all Irish samples and 44 of the 46 breeding programme bees were assigned to the M-lineage with a high degree of purity (Adapted from Henriques *et al*. 2018)

DISCUSSION

In this study, the year containing the highest count of colonies with late season mite loads ≥10% was 2014 with 77% of colonies below this crucial threshold (Delaplane and Hood 1997, Martin 1998). Slightly over 86% and 83% were below the same threshold for 2015 and 2016b respectively (Table 2). Unfortunately, despite this (and the presence of a considerable number of colonies that had a varroa load ≤2%), no beekeepers committed to leaving colonies untreated and record the outcome. This led to the failure of the programme despite it showing significant potential.

It was difficult to determine why this occurred. However, from talking to beekeepers, distrust of the sugar shake assay combined with a perceived high probability of the loss of a significant portion of colonies in a small apiary appears the most likely candidate. There were a small number of reports of low mite loads for colonies that were purportedly untreated for a number of years prior to this programme. Without proper data from the years prior to engagement with the programme, these cases remain anecdotal evidence of resistant colonies. However, they may be cases that are worthy of future investigation.

Although full approbation may not belong to this research, there is a sense that increasing numbers of beekeepers are trying their own method to improve varroa resistance, such as McMullan (2018). In addition, beekeepers in the south-east of the country established a small research apiary breeding for hygienic bees with the support of NUI Galway and NIHBS. An important factor helping to give impetus to these breeding programmes is perhaps the knowledge, and maybe a feeling of vindication, within NIHBS members that genetic proof of their claims about the persistence of *A. m. mellifera* and its comparative lack of hybridization had been obtained. That the spark for fresh research ignited from discussions with participants and presentations to BKAs underlines the importance of keeping stakeholders continuously informed about solid results, particularly in citizen science partnerships.

Any positive impetus imbued on stakeholders in scientific undertaken is all the more important when you consider that the attempted breeding programme described herein failed to achieve its original objective. This appears to have been as a consequence of a number of factors which can in all likelihood be traced back to how fiercely protective beekeepers are of their honey bee charges. We were lacking in beekeeping skills and underestimated the need to convince participants about the efficacy of the method. These

issues, compounded by a lack of protection for beekeepers in the event of colony loss underscored the failure.

Selection method rationale

There are a number of honey bee colony traits that are reported to increase resistance or tolerance of varroa (Harbo and Harris 1999b, Arechavaleta-Velasco and Guzman-Novoa 2001, Rosenkranz *et al.* 2010, Rinderer *et al.* 2010, Ardestani 2015) including drone entombment (Rath and Drescher 1990), hygienic and grooming behaviour (Spivak and Reuter 2001), forms of mite reproduction suppression (Harris and Harbo 1999a) and post-capping period (Bienefeld and Zautke 2007). However, there is considerable variability in the heritability value (h^2) of traits (Harbo and Harris 1999a) and the manner by which a particular trait is assessed, direct observation (Moretto *et al*. 1993) or proxy (Harbo and Harris 1999a), can have a considerable influence on the h^2 value produced. This makes it diffcult to optimise which resistance/tolerance traits to assay, paricularly in a citizen-science endeavour. Furthermore, for some traits, such as brood cell-size, experimental evidence suggests environmental and/or lineage-specific effects (Martin and Kryger 2002, Taylor *et al.* 2008, Seeley and Griffin 2011, Oddie *et al.* 2019). Additionally, selection assays which are difficult or time-consuming for beekeepers to undertake are likely to create a need for trained personnel to disseminate training and or to conduct the assays in place of the beekeeper.

As a consequence of the forgoing, resistance traits requiring specific evaluation by highly trained individuals were not chosen here. The sugar shaker assay (Lee *et al.* 2010, Macedo *et al.* 2002) to count the mites found on ~300 bees/100mL has fewer pre-requisites, takes only a few minutes to perform, is generally non-lethal to the bees and requires only one hive visit per count. Enhancing its appeal and utility amongst beekeepers is that it is readily interpreted and facilitates direct comparison discussions within the beekeeping community. A further consideration was that, as data collected by non-experts comes with an increased potential for data bias (Bird *et al.* 2014), the simplicity of the sugar shaker assay in combination with an assessment of colony characteristics already familiar to many beekeepers would mitigate against this effect.

When assays such as the freeze-kill brood for hygienic behaviour have a high h^2 value (0.65) (Harbo and Harris, 1999b) and the hygienic trait alone is shown to have a strong negative effect on mite reproduction rates, breeding programmes may ignore or lack the resources to

investigate the possibility of synergistic action between a variety of possible resistance and tolerance mechanisms. The health of a colony is strongly affected by its genetic diversity (Simone-Finstrom *et al.* 2016, Tarpy 2003, Tarpy *et al.* 2013) this is in part because the resultant variation in immunity to pathogens means that a single disease may not kill an entire colony but also because it creates an essential mix of forager types (Page *et al.* 1995, Page and Fondrk 1995, Pankiw and Page Jr 2000), ensuring adequate colony nutrition. Selection for only one particular resistance trait may also select for deleterious genes linked to it and this aspect is currently poorly understood. Consequently, any selection process, such as the sugar shaker assay, that focusses on the final outcome rather than the trait behind it has a greater likelihood of retaining genetic diversity, producing healthier colonies, and avoiding losing the epistatic effects of multiple genes (Behrens *et al.* 2011) that may be required to each contribute their part to forms of tolerance as yet unelucidated.

Incoming data from beekeepers

Participants agreed to conduct testing on two occasions during the year. Multiple mite counts (4+) as each colony developed through the year, particularly during the period of greatest colony growth, such as recommended for the BeeBreed breeding values assessment (Länderinstitut für Bienenkunde, Bee institute, Hohen Neuendorf, Germany) were sought but advised against by NIHBS on the basis that the rigours of multiple testing would be likely to result in a poor uptake by beekeepers. The agreed test times were as soon after overwintering as weather conditions allowed (this could vary considerably around the country) and immediately before the Autumn acaricide application to allow for non-treatment where indicated. Unfortunately, in 2014 and 2015 only one test report was received per apiary. The reason for this was discussed up with a small selection of the beekeepers and a lack of available time during apiary visits combined with insufficient periods of good weather were given as the primary reasons for non-responses. Ireland has a short season of nectar flow relative to mainland Europe, combined with low annual sunshine values and high rainfall. Consequently, time to conduct ordinary hive manipulations is at a premium and non-essential work such as surveys understandably does not get preference.

The dates of the single tests that were submitted for 2014 and 2015 ranged between May to October although most were in the Autumn. The weather in Ireland, particularly in the north west can be very wet into May and this prevented some beekeepers from opening hives

during this period. This meant that for some beekeepers the time of the first test probably coincided with the swarming season which usually begins in early to mid-May in Ireland, thereby further reducing the opportunities to conduct assays.

In addition to only one test return in each year, the continuity of data on individual colonies was poor (Table 3). It is possible that in short term citizen science, adherence to a rigorous test protocol, however simplified, may not yield significantly useful data and that a research facility or dedicated team of field researchers may be necessary to achieve sufficient data in the honey bee arena.

The values obtained from the mainly Autumnal varroa counts of 2014 and 2015 do not differ significantly from the 2016 Autumn data with mean mite counts per 300 bees of 7.69, 4.55 and 5.13 respectively (Table 1). Interestingly the range of mean values for colonies with mite loads of <2% and zero are also not significantly different between the three years. The relatively small value ranges give some validity to the precision of the testing method when used by citizen scientists. At the Autumnal time of testing in 2014, 2015 and 2016b, there were between 43.1% and 51.9%% of colonies expressing ≤2% mite load indicating that there were sufficient colonies available for beekeepers to attempt withholding acaricide treatment per the breeding programme guidelines. All colonies were subsequently treated with acaricides which was an unexpected outcome, since the aims and rationale of the programme had been agreed beforehand and circulated freely.

An unwillingness of beekeepers to accept the validity of the testing method and act to discontinue treatment on cases with <2% load became evident during the data collection. It could be that the amount of time needed to engage with participants to optimise their faith in the methodology was underestimated. The same could be said about the recording of the additional colony characteristics. These were to determine if particular colony characteristics had greater association with untreated colonies and to investigate if the brood proportions had a significant effect on the mite loads for those colonies. Some participants indicated dissatisfaction with the time taken to complete these sections which may have been led to a reduction in re-recruitment for subsequent tests. It was thought that the time taken by participants to expedite the data recording would reduce with experience. Prior use of a larger focus group to pre-assess protocol timings may have revealed that we needed to devote a greater amount of time to participant training rather than recruitment.

The lack of continuous data on untreated colonies meant the supplementary data were not investigated further in this study. In future, simplification of the participants' involvement to only require the sugar shaker assay may have provided more continuous data. It is worth repeating though that this would only be feasible if greater faith in the sugar-shake assay could persuade participants to engage with non-treatment. The information on characteristics could be sought after a period of non-treatment had been recorded.

Some participants indicated that the second test point in August came at a time when they were too busy with honey extraction. Lack of time was cited as the main reason why multiple assessment times during the year were thought to be inoperable, including the spring assessment. Time will always be at a premium for beekeepers, especially commercial ones, therefore in any similar future programme, keeping participants motivated is likely to form a key factor to encourage the additional work required.

Breeding programmes

In local breeding programmes such as organised by a BKA or undertaken by sole beekeepers there may be a knowledge gap in understanding between the concepts of rearing and of selective breeding (Uzunov *et al.* 2017) which would need to be addressed at an early stage of the development of the programme.

Existing varroa tolerance breeding programmes such as the SMARTBEES consortium and the Arbeitsgemeinschaft toleranzzucht (AGT, www.toleranzzucht.de) programme in Germany initially select for hygienic behaviour using either the Pin kill brood (PKB) or Freeze kill brood (FKB) assay (Uzunov *et al*. 2015) to identify a possible colony exhibiting a varroa specific hygiene (VSH) resistance mechanism. There is not universal agreement on the use of the FKB assay in assessing colonies for hygienic behaviour (Leclercq *et al.* 2018a). The accuracy and fidelity between colonies of the PKB assay can be biased by the chosen pin diameter (Leclercq *et al.* 2018b) which must remain uniform for all colonies. This characteristic of the assay may also bias data comparisons between test facilities. The SMARTBEES programme runs training days for participants to ensure each one has had sufficient instruction in the methodology. These arguments cloud the view of the best way forward when designing a novel breeding programme. For many beekeepers maintaining colony characteristics favourable to easier bee handling and improved honey production beekeeping is of utmost importance.

Therefore, since varroa tolerant honey bees with undesirable characteristics may not be desirable for breeding (Nicodemo *et al.* 2013), both SMARTBEES and AGT record a range of colony characteristics to inform decision making and roughly the same system was incorporated into this programme.

Although not breeding programmes for varroa tolerance in the strict sense, numerous investigations into colony survival in the presence of varroa have taken place such as "James Bond" or "Live and let die" type experiments. These are named after the film of the same name and involve leaving a number of colonies unmanaged and untreated to observe the survivorship over time. Evaluating survival of untreated stock (Le Conte *et al.* 2007, De Guzman *et al.* 2001, Rinderer *et al.* 2010) and longevity of free-living colonies may elucidate new tolerance mechanisms and genotypes for future inclusion in breeding selection. It appears that from a reasonable stock size, with genetically diverse sources including non-native subspecies and commercial hybrid strains, there may be sufficient colonies with characteristics suitable to surviving varroa infestation (Fries *et al.* 2006). Although Fries *et al.* (2006) found that their survivor colonies displayed increased swarming behaviour, no link between this behaviour and a reduced varroa load was seen, indicating that it is not the tolerance mechanism, at least not in the cold Nordic environment. A possible link between hybridisation and varroa tolerance is also seen in wild or feral colonies (Seeley 2007) which raises the question; Can genetically isolated *A. m. mellifera* naturally survive varroa or is the hybridisation reported in these studies a product of the beekeeping environment in those countries?

In the early development of any future breeding programme for Ireland, a course of action may be to select beekeepers who have some experience (but do not need to be highly experienced), that understand the concepts of selective breeding and have the support of other beekeepers, freeing them to proceed with the confidence that failures and losses will be recompensed in kind.

Selective breeding in Ireland

Articles in Beekeeping Association (BKA) periodicals, winter talks and practical demonstrations on the methods of queen rearing are myriad (Pers. obs.) and allude to the subject being difficult to grasp for many beekeepers. Despite access to the abundance of information within the Irish beekeeping community, the evolution of basic queen rearing

operations into a selective breeding programme involving the evaluation and selection of donor colonies combined with a controlled evaluation of the colonies produced, seems to be a rare event.

Beekeeper numbers in Ireland are difficult to determine since not all are members of BKAs but a number of between 2500 and 6000 seems reasonable through a combination of official figures and anecdotal evidence. The beekeeper demographic in Ireland is primarily hobbyist, producing the lowest volumes of honey and numbers of queens of any European country (Chauzat *et al.* 2013). The smaller number of commercial beekeeping concerns, which tend to practice early queen replacement, has possibly resulted in a dearth of beekeepers with skills in the selective breeding of queens.

The Galtee Bee Breeding Group (GBBG) based in County Tipperary is one of the longest established breeding programmes on the island of Ireland. Their primary aim is to produce local native, black honey bee colonies where the most sought-after characteristics are carefully assessed and recorded in studbooks. The group used wing morphometry to assess for M-lineage purity in order to use only local *A. m. mellifera* and included artificial insemination techniques to assist the continuation of the colony characteristics between generations.

Our understanding was, that as our programme was being undertaken, there was no formal breeding in Ireland for resistance to *V. destructor*.

It may have always been beyond the remit of this research effort to fully instigate an all-Ireland breeding programme without solid pre-existing credentials in both honey bee genetics and, more importantly, beekeeping. However, encouraged by the efforts of our programme and with the understanding that resistance characteristics are heritable in other environments with open-mated queens (Harbo and Harris 2001), a research apiary to assess the heritability of hygienic behaviour in an open-mating environment in Ireland was established in the South East of the country in collaboration with NIHBS and NUI Galway. If this has successes which are repeatable it may pave the way for further programmes.

Necessity of using local ecotypes for selection purposes.

The evolution of *Apis* species has produced a eusocial insect honed by the selection pressures placed on it by its environment. For example, bee body size and pigmentation along with body hair length and density are considered dependant on mean annual temperatures and sunlight

levels with variations in both morphological and physiological phenotypes existing in honey bees from different environments (Bouga *et al.* 2005, Parker *et al.* 2010). Reduced foraging effort relative to local bees can occur in imported conspecifics (Alqarni 2006). Surviving the wet and cold of winter and the efficacy of the spring revival by the colony are perhaps the two of the greatest challenges faced by colonies in the northern hemisphere. It is worth considering whether the increasing availability of environmental forage in spring or prior increase in colony size occur, driven by available stores (possibly favouring greater overwintering ability), is the dominant driver of colony build-up. If the overwintering ability and brood size (Hatjina *et al.* 2014) of local ecotypes coordinate more closely with the availability of forage than other conspecifics, this could confer them with a competitive advantage. Positive outcomes to colony activities such as the reproductive strategies of drone production and swarming are highly dependent on available energy (Starr 2006) which may, in turn, depend on how available forage is utilised. The social immunity of honey bees is built on colony characteristics such as hygienic behaviour (Spivak and Gilliam 1998a, Spivak and Gilliam 1998b, Momot and Rothenbuhler 1971), whose efficacy can be affected by environmental conditions (Momot and Rothenbuhler 1971), and even to integumental chemicals (Del Piccolo *et al.* 2010, Frey *et al.* 2013) whose precursor molecules may be derived from specific forage. If an indigenous ecotype has adaptations specific to the local environment (Bouga *et al.* 2005, Parker *et al.* 2010, Costa *et al.* 2012, Büchler *et al.* 2014) the introduction of conspecifics whose persistence is propped up by anthropogenic means (e.g. regular re-importation in the face of losses, subsidised feeding or treatment for diseases) has the potential to overwhelm the ecotypic genotype, remove locally adapted genes and endanger the long-term survival of local honey bees. Because of the greater pollination effort (Al-Ghamdi *et al.* 2017) and colony longevity afforded by local ecotypes (Costa *et al.* 2012, Büchler *et al.* 2014) the consequences of ecotype loss can potentially include a negative impact on national food security by increasing reliance on the importation of non-native queens.

Preserving the native ecotype

The most effective way to prevent further introgression from introduced conspecifics is to remove them from the environment thus removing the exogenous material (Amador *et al.* 2013). Failing this, the logical step would be to stop their continued reintroduction. In Ireland,

and the United Kingdom, there are moves to obtain a ban on the continued importation of conspecifics in an effort to mitigate introgression and the potential loss of native ecotypes (NIHBS, 2021). In the European Union honey bees are considered domesticated livestock and can be sold between member states under its Common Agricultural Policy. Consequently, any legal effort to attempt to end their movement between member states is likely to be considered as undermining one of the fundamental principles of the EU and may encounter considerable bureaucratic resistance.

Currently, not all of Ireland's beekeepers use *A. m. mellifera*. While some may have made a conscious decision to keep a conspecific or commercial hybrid, it is likely that many do not know the genetic provenance of their bees and/or may be unaware of the probable consequences surrounding the loss of native ecotypes. Informing the beekeeping community about the native ecotype and how to help it persist is therefore essential and is the central work of NIHBS.

As long as free-mating of queens remains the norm in Ireland, where the placement and movement of apiaries is unregulated, there will always be a hybridization risk for *A. m. mellifera*. Checks on beekeeping social media sites, where these issues are often raised, reveal that keepers of conspecifics and commercial hybrids feel there is a mirror problem for their queens. Government interest in apiaries extends only to the registration of those where primary production of honey is intended, therefore the onus is on beekeepers alone to initiate a solution to the issue. A viable option may include shared information on the location (these may be rough coordinates since hive theft is a serious consideration) of apiaries of the different breeds of bee as this would allow the positioning of mating stations away from the unwanted exogenous material.

Future breeding programme

It appears that for a nationwide breeding programme to succeed some form of compensation for losses is required for beekeepers who experience the death of a colony or queen. A possible cost-effective compensation method which provides free, altruistic, replacement nuclei between groups of participants or from their BKA would encourage the greater, but controlled, risk taking required to start the selection process. Provided sufficient beekeepers agree to participate, the wide geographic spread of BKAs could provide a nationwide support network for a breeding programme, reducing the risk to individual beekeepers.

A proposed organisational flow-plan for a breeding programme to be successful in Ireland is illustrated in Figure 4. BKA member beekeepers would test for varroa infestation at two or more time points and donate putative resistant colonies or mated queens to a centralised test apiary (established by the BKA) for ongoing assessment and accurate record keeping of varroa counts and colony characteristics per the GBBG protocol. Virgin queens raised from the most resistant colonies of the test apiary along with similar queens from other BKAs are mated at an isolated mating station populated by drone production colonies produced from the test apiaries. This plan allows for support from within their local breeding group or BKA, of participating breeders and those who wish to adhere to the original programme of testing and re-queening within their own apiary.

Each BKA test apiary populated with putative resistant colonies donated from the membership, plus an isolated mating station shared between several local associations could facilitate cross breeding between a diversity of resistant genotypes provided issues of lineage introgression were addressed.

The model would require the establishment of isolated mating locations and these same locations may suffice to allow controlled mating of pure *A. m. mellifera* from within the cohort of beekeepers that wish to avoid introgression from other sub-species. In deciding on locations for mating apiaries care needs to be taken to investigate the surrounding region and to liaise with all beekeepers, regardless of the provenance of their bees.

Islands off Ireland where the prevailing wind blows away from the mainland have been considered and their efficacy investigated (Morgan 2018) however even an island location does not guarantee 100% isolation (Kraus 2005), making knowledge of the source location of genetic material imperative to success.

If resistance/tolerance to V*arroa destructor* is the primary breeding aim that, along with the potential of reduced hive manipulation time and operating costs, may generate sufficient motivation and conviction from beekeepers. However, any breeding programme that can also create a revenue stream would help attract a sufficient source hives to ensure genetic diversity. For instance, queens certified as varroa tolerant, from properly evaluated colonies should attain a premium value. Although open-mated queens raised from a tolerant colony may continue to provide the desired tolerance (Harbo and Harris 2001), unless the tolerance stems from a characteristic that is highly heritable (Harbo and Harris 1999a and 1999b) performance may not meet standards expected by buyers and any resultant inefficacy could

undermine the revenue stream. In mitigation, controlled mating stations using a mixture of the best performing colonies both for drone production and to generate queens could be used to aim for an optimal combination of tolerance to varroa along with desired characteristics such as easy manageability and good honey production. Sufficient genetic diversity is possible by maintaining a studbook for each queen lineage and encouraging breeders to utilise a number of mating stations.

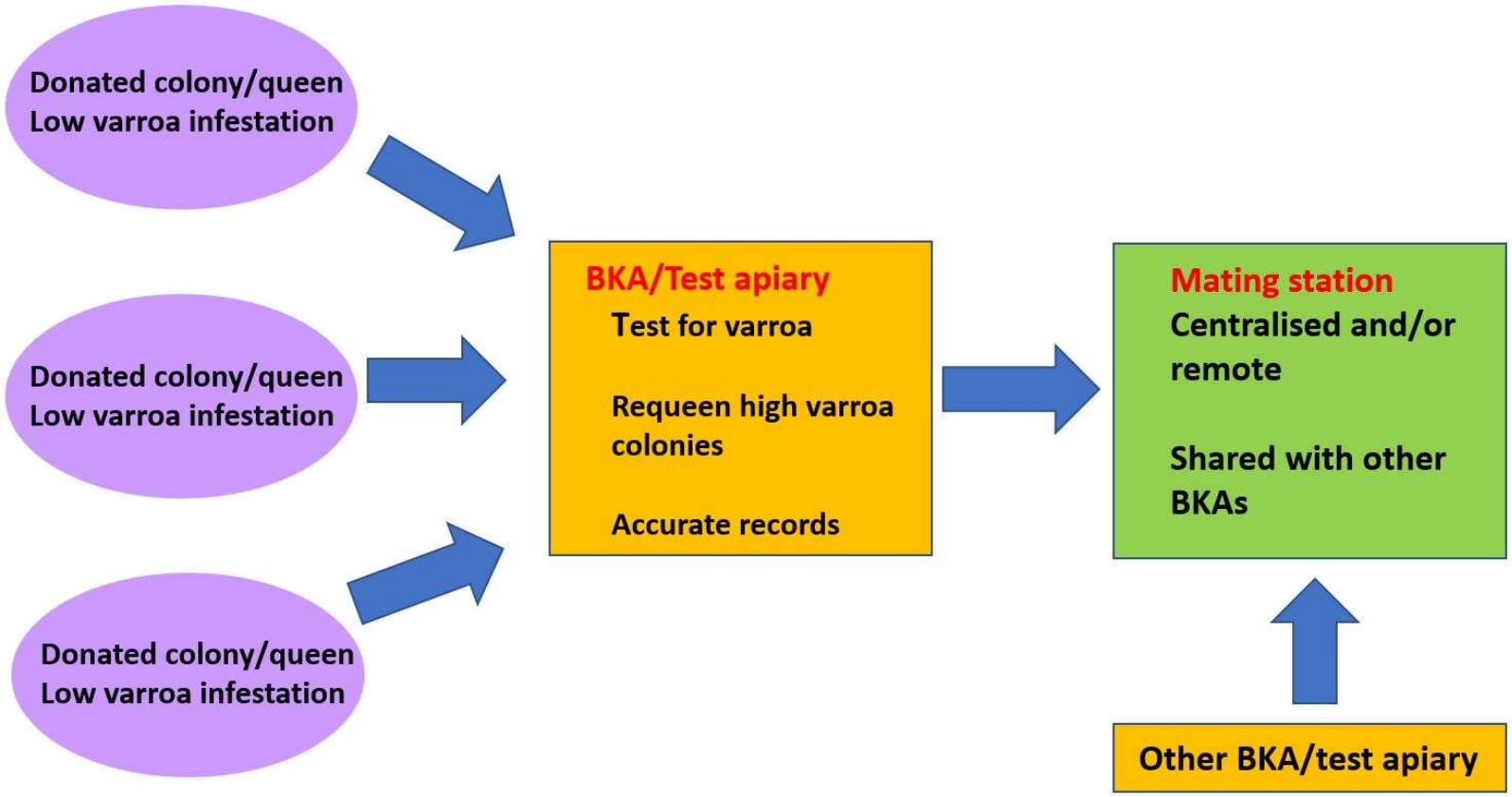


Figure 4: Organisational flow-chart for a suggested breeding programme to produce resistant colonies. BKA member beekeepers (purple), centralised test apiary (orange), other BKAs/test apairies (orange box/blue arrows), isolated mating station (green)

If local programmes prove successful and a nationwide breeding programme was to be implemented in the future, it will require the appointment of full-time trainers and administrative staff to aid researchers. Trained individuals, who in turn train other beekeepers how to enact the findings of scientific research can also act as fieldwork operatives. Research continues to investigate the mechanisms of resistance to or tolerance of *V. destructor* and where a fresh mechanism has been identified this would need to be communicated to beekeepers as soon as practicable by trainers in order to make a real-world difference.

Amelioration of *Varroa destructor* effects

In Ireland, as elsewhere, the continued reliance on chemicals is an expected consequence of the initial treatments which were considered necessary to ameliorate widespread catastrophic losses when varroa was first discovered in *A. mellifera*. Since this time, integrated pest management (IPM) which includes biotechnical controls such as drone brood trapping and brood breaks (Calis *et al.* 1999, Delaplane *et al.* 2005, Dietemann *et al.* 2012) are increasingly being put forward as a way to reduce or even remove the reliance on acaricides. Techniques such as the brood controls above are the product of using an understanding of the biology of varroa to negatively influence its reproductive success with the aim of preventing infestation from reaching levels that overwhelm the colony and are relatively cost-free apart from the time spent manipulating the colony. Since chemicals are not used, they retain the concept of pure honey that is highly prized by beekeepers and consumers alike.

Despite these positive elements, lack of widespread knowledge about the methods, doubts about efficacy or the increased time element of their application relative to chemical treatments, appears to have resulted in a slow uptake of these forms of biotechnical defence within the Irish beekeeping community. Our results (Appendix BP VIII) indicate that chemical treatment appears to be used prophylactically throughout an apiary either/or in the spring or directly after Autumnal honey harvest without much consideration for or to prior assessment of the infestation levels. This preventative form of treatment mirrors other forms of animal husbandry in Ireland and elsewhere (Martin *et al.* 2020) which treat to prevent rather than cure disease but run the risk of producing resistant pathogens (Milani 1999, Martin 2015, Rinkevich 2020). The overall picture of non-assessment was a surprise given that the majority

of beekeepers are part of BKAs which regularly disseminate current topics on beekeeping including biotechnical varroa control.

Where varroa assessments are carried out, daily natural mite drops are commonly used as a means of assessing infestation levels. These require an apiary visit to set the bottom boards and a second a few days later to count the mites found on it. Results from this method can be ambiguous for a number of reasons. Live mites may fall and then return to the colony unless a film of oil or petroleum jelly is used to prevent this happening. Dead mites may result from senescence however this number can be added to by auto and allo-grooming by the bees (Boecking and Spivak 1999). Where grooming occurs, determining the number of normal fatalities from mites killed by grooming or post mortem scavenging by other hive inhabitants is not straightforward. It is possible to misinterpret damage caused by wax moth larvae and other invertebrate predators and scavengers to naturally dead mites as grooming damage (Bienefeld *et al.* 1999). Even where operators may be practiced and proficient enough to determine the disparity in damage types it would be difficult to accurately compare results with neighbouring beekeepers and obtain a wider picture of local or regional infestation unless all follow identical protocols. To accurately decipher marks on the mite corpses also takes time and requires a level of magnification which is difficult to complete in the field.

Apart from the possibility of mite adaptation to acaracides, some of these chemicals also have implications for the health of both the consumer and the beekeeper, it follows that selection for forms of natural resistance or tolerance to *V. destructor* creates a more sustainable future for honey production.

Genotyping

Sixty percent of the Irish mitotype variants were identical to two Netherland samples. This equates well with the known widespread losses Ireland's honey bee population underwent in the early part of 20th century as a consequence of Isle of Wight disease (Rennie *et al.* 1921) and the subsequent importation of quantities of skeps of Dutch bees.

The comparative lack of C-lineage introgression relative to other European populations which was confirmed in the three methods may be due to the relatively low volume of honey bee imports into Ireland. Similarly, it may reflect a general preference amongst Irish beekeepers for the dark bee and a hobbyist profile that gives preference to the domestic production of queens. Evidence of M-lineage was found in all bees mitochondrially genotyped. Although

this was a very small sample size it raises the possibility of assortative queen mating occurring in Ireland through mate selection by the *A. m. mellifera* queens or other forms of reproductive isolation between sympatric sub-species (Oleksa *et al.* 2013) such as spatial (Koeniger *et al.* 1989, Rowell *et al.* 1992) or temporal (Jaffe *et al.* 2009) distributions of drones within drone congregation areas (DCAs). These mating dynamics, which favour conspecific mating at the sub-species level, could have a positive effect on the persistence of the M-lineage in Ireland as well as influencing geneflow for heritable resistance traits and local adaptations. It is important to note however that since honey bees are highly polyandrous, those tested in this study only represent a small proportion of the drone diversity in their colonies and more intensive testing may be needed to reveal a clearer picture of C-lineage introgression levels.

Whatever the reason or mechanism behind the existence of pure *A. m. mellifera* in Ireland it bolsters the need to produce a breeding programme which helps protect it from losses caused by *V. destructor*.

CONCLUSION

Overall, the mite percentage data collected showed considerable variation in either the potential for colonies to respond to treatments or in the application of treatments by beekeepers. While the lack of continuous untreated colony data was unfortunate, there were a number of untreated colonies with a varroa load of <2% and these have the potential to form the basis of future efforts to breed varroa-resistant colonies.

The understanding and support of keepers of all types of honey bees in Ireland is needed for any selective breeding programme to be successful. Thankfully, the concept of establishing a breeding programme to bolster any resistance to varroa that does exist in Ireland's honey bees has itself received a boost from the evidence of the existence of pure *A. m. mellifera* within the country.

The work here centred only on bees bred by members of NIHBS, consequently genotyping needs to be conducted extensively outside this group to gain a clear countrywide picture. The combination of pure *A. m. mellifera* and a breeding programme against the ravages of *V. destructor* has the potential to move Ireland into the forefront of honey bee research and conservation in Europe.

REFERENCES

Alexander, D. H., Novembre, J. and Lange, K. (2009) 'Fast model-based estimation of ancestry in unrelated individuals', *Genome Research,* 19(9), 1655-1664.

Al-Ghamdi, A. A., Adgaba, N., Tadesse, Y., Getachew, A. and Al-Maktary, A. A. (2017) 'Comparative study on the dynamics and performances of *Apis mellifera jemenitica* and imported hybrid honeybee colonies in southwestern Saudi Arabia', *Saudi Journal of Biological Sciences,* 24(5), 1086-1093.

Alice Pinto, M., Henriques, D., Chavez-Galarza, J., Kryger, P., Garnery, L., van der Zee, R., Dahle, B., Soland-Reckeweg, G., de la Rua, P., Dall'Olio, R., Carreck, N. L. and Johnston, J. S. (2014) 'Genetic integrity of the Dark European honey bee (*Apis mellifera mellifera*) from protected populations: a genome-wide assessment using SNPs and mtDNA sequence data', *Journal of Apicultural Research,* 53(2), 269-278.

Alqarni, A. (2006) 'Tolerance of summer temperature in imported and indigenous honeybee *Apis mellifera* L. races in central Saudi Arabia', *Saudi Journal of Biological Sciences,* 13.

Amador, C., Fernández, J. and Meuwissen, T. H. (2013) 'Advantages of using molecular coancestry in the removal of introgressed genetic material', *Genetics, selection, evolution*: GSE, 45(1), 13-13.

Anderson, D. L. and Trueman, J. W. H. (2000) '*Varroa jacobsoni* (Acari : Varroidae) is more than one species', *Experimental and Applied Acarology,* 24(3), 165-189.

Behrens, D., Huang, Q., Gessner, C., Rosenkranz, P., Frey, E., Locke, B., Moritz, R. F. A. and Kraus, F. B. (2011) 'Three QTL in the honey bee *Apis mellifera* L. suppress reproduction of the parasitic mite varroa destructor', *Ecology and Evolution,* 1(4).

Bienefeld, K., Zautke, F., Pronin, D. and Mazeed, A. (1999) 'Recording the proportion of damaged *Varroa jacobsoni* Oud. in the debris of honey bee colonies (*Apis mellifera*)', *Apidologie,* 30(4), 249-256.

Bird, T. J., Bates, A. E., Lefcheck, J. S., Hill, N. A., Thomson, R. J., Edgar, G. J., Stuart-Smith, R. D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J. F., Pecl, G. T., Barrett, N. and Frusher, S. (2014) 'Statistical solutions for error and bias in global citizen science datasets', *Biological Conservation,* 173, 144-154.

Blacquière, T. and Panziera, D. (2018) 'A plea for use of honey bees' natural resilience in beekeeping', *Bee World,* 95(2), 34-38.

Boecking, O. and Spivak, M. (1999) 'Behavioral defenses of honey bees against *Varroa jacobsoni* Oud', *Apidologie,* 30(2-3), 141-158.

Boerjan, M., Kok, H.S., Kok and Touw, E. (2018) Natural selection of varroa tolerant *Apis mellifera mellifera* "black bees" in the Netherlands, Finland S.I.C.A.M.M. conference abstract

Bouga, M., Kilias, G., Harizanis, P. C., Papasotiropoulos, V. and Alahiotis, S. (2005) 'Allozyme variability and phylogenetic relationships in honey bee (Hymenoptera: Apidae: *Apis mellifera*) populations from Greece and Cyprus', *Biochemical Genetics,* 43(9-10), 471-483.

Browne, K., McCormack, G.P. (2015) 'The search for tolerance to varroa in Irish honey bees' *The Four Seasons, The newsletter of the Native Irish Honey Bee Society*, 56, 28-29

Büchler, R., Berg, S. and Le Conte, Y. (2010) 'Breeding for resistance to *Varroa destructor* in Europe', *Apidologie,* 41(3), 393-408.

Büchler, R., Costa, C., Hatjina, F., Andonov, S., Meixner, M. D., Conte, Y. L., Uzunov, A., Berg, S., Bienkowska, M., Bouga, M., Drazic, M., Dyrba, W., Kryger, P., Panasiuk, B., Pechhacker, H., Petrov, P., Kezić, N., Korpela, S. and Wilde, J. (2014) 'The influence of genetic origin and its interaction with environmental effects on the survival of *Apis mellifera* L. colonies in Europe', *Journal of Apicultural Research,* 53(2), 205-214.

Calis, J. N. M., Boot, W. J., Beetsma, J., van den Eijnde, J. H. P. M., de Ruijter, A. and van der Steen, J. J. M. (1999) 'Effective biotechnical control of varroa: applying knowledge on brood cell invasion to trap honey bee parasites in drone brood', *Journal of Apicultural Research,* 38(1-2), 49-61.

Chauzat, M. P., Cauquil, L., Roy, L., Franco, S., Hendrikx, P. and Ribiere-Chabert, M. (2013) 'Demographics of the European Apicultural Industry', *Plos One,* 8(11).

Clement, M., Posada, D. and Crandall, K. A. (2000) 'TCS: a computer program to estimate gene genealogies', *Molecular Ecology,* 9(10), 1657-1659.

Cornuet, J. M., Garnery, L. and Solignac, M. (1991) 'Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA', *Genetics,* 128(2), 393-403.

Costa, C., Lodesani, M. and Bienefeld, K. (2012) 'Differences in colony phenotypes across different origins and locations: evidence for genotype by environment interactions in the Italian honeybee (*Apis mellifera ligustica*)?', *Apidologie,* 43(6), 634-642.

Dejong, D., Roma, D. D. and Goncalves, L. S. (1982) 'A comparative analysis of shaking solutions for the detection of *Varroa jacobsoni* on adult honeybees.1', *Apidologie,* 13(3), 297-306.

Del Piccolo, F., Nazzi, F., Della Vedova, G. and Milani, N. (2010) 'Selection of *Apis mellifera* workers by the parasitic mite *Varroa destructor* using host cuticular hydrocarbons', *Parasitology,* 137(6), 967-973.

Delaplane, K. S., Berry, J. A., Skinner, J. A., Parkman, J. P. and Hood, W. M. (2005) 'Integrated pest management against *Varroa destructor* reduces colony mite levels and delays treatment threshold', *Journal of Apicultural Research,* 44(4), 157-162.

Delaplane, K. S. and Hood, W. M. (1997) 'Effects of delayed acaricide treatment in honey bee colonies parasitized by *Varroa jacobsoni* and a late-season treatment threshold for the south-eastern USA', *Journal of Apicultural Research,* 36(3-4), 125-132.

Delaplane, K. S. and Hood, W. M. (1999) 'Economic threshold for *Varroa jacobsoni* Oud. in the southeastern USA', *Apidologie,* 30(5), 383-395.

De la Rua, P., Jaffe, R., Dall'Olio, R., Munoz, I. and Serrano, J. (2009) 'Biodiversity, conservation and current threats to European honeybees', *Apidologie,* 40(3), 263-284.

Dietemann, V., Pflugfelder, J., Anderson, D., Charriere, J. D., Chejanovsky, N., Dainat, B., de Miranda, J., Delaplane, K., Dillier, F. X., Fuch, S., Gallmann, P., Gauthier, L., Imdorf, A., Koeniger, N., Kralj, J., Meikle, W., Pettis, J., Rosenkranz, P., Sammataro, D., Smith, D., Yanez, O. and Neumann, P. (2012) '*Varroa destructor*: research avenues towards sustainable control', *Journal of Apicultural Research,* 51(1), 125-132.

Ellis, A. M., Hayes, G. W. and Ellis, J. D. (2009) 'The efficacy of dusting honey bee colonies with powdered sugar to reduce varroa mite populations', *Journal of Apicultural Research,* 48(1), 72-76.

Finnoff, D., Shogren, J. F., Leung, B. and Lodge, D. (2007) 'Take a risk: Preferring prevention over control of biological invaders', *Ecological Economics,* 62(2), 216-222.

Frey, E., Odemer, R., Blum, T. and Rosenkranz, P. (2013) 'Activation and interruption of the reproduction of *Varroa destructor* is triggered by host signals (Apis mellifera)', *Journal of Invertebrate Pathology,* 113(1), 56-62.

Fries, I., Imdorf, A. and Rosenkranz, P. (2006) 'Survival of mite infested (*Varroa destructor)* honey bee (*Apis mellifera*) colonies in a Nordic climate', *Apidologie,* 37(5), 564-570.

Garnery, L., Cornuet, J. M. and Solignac, M. (1992) 'Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis', *Molecular Ecology,* 1(3), 145-154.

Garnery, L., Franck, P., Baudry, E., Vautrin, D., Cornuet, J.-M. and Solignac, M. (1998a) 'Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). I. Mitochondrial DNA', *Genetics Selection Evolution,* 30, S31-S47.

Gregorc, A., Knight, P. R. and Adamczyk, J. (2017) 'Powdered sugar shake to monitor and oxalic acid treatments to control varroa mites (*Varroa destructor* Anderson and Trueman) in honey bee (*Apis mellifera*) colonies', *Journal of Apicultural Research,* 56(1), 71-75.

Harbo, J. R. and Harris, J. W. (1999a) 'Heritability in honey bees (Hymenoptera : Apidae) of characteristics associated with resistance to *Varroa jacobsoni* (Mesostigmata : Varroidae)', *Journal of Economic Entomology,* 92(2), 261-265.

Harbo, J. R. and Harris, J. W. (1999b) 'Selecting honey bees for resistance to *Varroa jacobsoni*', *Apidologie,* 30(2-3), 183-196.

Harbo, J. R. and Harris, J. W. (2001) 'Resistance to *Varroa destructor* (Mesostigmata : Varroidae) when mite-resistant queen honey bees (Hymenoptera : Apidae) were free-mated with unselected drones', *Journal of Economic Entomology,* 94(6), 1319-1323.

Harris, J., Rinderer, T., Kuznetsov, V., Danka, R., Delatte, G., De Guzman, L. and Villa, J. (2002) 'Imported Russian honey bees: Quarantine and initial selection for varroa resistance', *American Bee Journal,* 142(8), 591-596.

Hassett, J., Browne, K. A., McCormack, G. P., Moore, E., Soland, G., Geary, M. and Native Irish Honey Bee, S. (2018) 'A significant pure population of the dark European honey bee (*Apis mellifera mellifera*) remains in Ireland', *Journal of Apicultural Research,* 57(3), 337-350.

Hatjina, F., Costa, C., Buechler, R., Uzunov, A., Drazic, M., Filipi, J., Charistos, L., Ruottinen, L., Andonov, S., Meixner, M. D., Bienkowska, M., Dariusz, G., Panasiuk, B., Le Conte, Y., Wilde, J., Berg, S., Bouga, M., Dyrba, W., Kiprijanovska, H., Korpela, S., Kryger, P., Lodesani, M., Pechhacker, H., Petrov, P. and Kezic, N. (2014) 'Population dynamics of European honey bee genotypes under different environmental conditions', *Journal of Apicultural Research,* 53(2), 233-247.

Henriques, D., Browne, K. A., Barnett, M. W., Parejo, M., Kryger, P., Freeman, T. C., Munoz, I., Garnery, L., Highet, F., Jonhston, J. S., McCormack, G. P. and Pinto, M. A. (2018) 'High sample throughput genotyping for estimating C-lineage introgression in the dark honeybee: an accurate and cost-effective SNP-based tool', *Scientific Reports,* 8.

Ibrahim, A. and Spivak, M. (2006) 'The relationship between hygienic behaviour and suppression of mite reproduction as honey bee (*Apis mellifera*) mechanisms of resistance to *Varroa destructor*' *Apidologie,* 37(1), 31.

Jaffe, R., Dietemann, V., Crewe, R. M. and Moritz, R. F. A. (2009) 'Temporal variation in the genetic structure of a drone congregation area: an insight into the population dynamics of wild African honeybees (*Apis mellifera scutellata*)', *Molecular Ecology,* 18(7), 1511-1522.

Jensen, A. B., Palmer, K. A., Boomsma, J. J. and Pedersen, B. V. (2005) 'Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honeybee, *Apis mellifera mellifera*, in northwest Europe', *Molecular Ecology,* 14(1), 93-106.

Koeniger, G., Koeniger, N., Pechhacker, H., Ruttner, F. and Berg, S. (1989) Assortative mating in a mixed population of European honeybees*, Apis mellifera ligustica* and *Apis mellifera carnica*. Insectes Sociaux 36, 129-138

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. and Mayrose, I. (2015) 'Clumpak: a program for identifying clustering modes and packaging population structure inferences across K', *Molecular Ecology Resources,* 15(5), 1179-1191.

Kraus, F.B., (2005) Requirements for local population conservation and breeding, *Beekeeping and conserving biodiversity of honey bees* pp 105, Eds. Lodesani M. and Costa C., Northern Bee Books, Hebden Bridge, UK

Le Conte, Y., De Vaublanc, G., Crauser, D., Jeanne, F., Rousselle, J.-C. and Becard, J.-M. (2007) 'Honey bee colonies that have survived *Varroa destructor*', *Apidologie,* 38(6), 566-572.

Leclercq, G., Blacquière, T., Gengler, N. and Francis, F. (2018a) 'Hygienic removal of freeze-killed brood does not predict varroa-resistance traits in unselected stocks', *Journal of Apicultural Research,* 57(2), 292-299.

Leclercq, G., Francis, F., Gengler, N. and Blacquière, T. (2018b) 'Bioassays to quantify hygienic behavior in honey bee (*Apis Mellifera* L.) colonies: A review', *Journal of Apicultural Research,* 57(5), 663-673.

Lee, K. V., Moon, R. D., Burkness, E. C., Hutchison, W. D. and Spivak, M. (2010) 'Practical sampling plans for *Varroa destructor* (Acari: Varroidae) in *Apis mellifera* (Hymenoptera: Apidae) colonies and apiaries', *Journal of Economic Entomology,* 103(4), 1039-1050.

Macedo, P. A., Wu, J. and Ellis, M. D. (2002) 'Using inert dusts to detect and assess varroa infestations in honey bee colonies', *Journal of Apicultural Research,* 41(1-2), 3-7.

Martin, H., Manzanilla, E. G., More, S. J., O'Neill, L., Bradford, L., Carty, C. I., Collins, Á. B. and McAloon, C. G. (2020) 'Current antimicrobial use in farm animals in the Republic of Ireland', *Irish Veterinary Journal,* 73(1), 11.

Martin, S. (1998) 'A population model for the ectoparasitic mite *Varroa jacobsoni* in honey bee (Apis mellifera) colonies', *Ecological Modelling,* 109(3), 267-281.

Martin, S. (2015) 'Acaricide (pyrethroid) resistance in *Varroa destructor*', *Bee World,* 85, 67-69.

Martin, S. J. and Kryger, P. (2002) 'Reproduction of *Varroa destructor* in South African honey bees: does cell space influence varroa male survivorship?', *Apidologie,* 33(1), 51-61.

Milani, N. (1999) 'The resistance of *Varroa jacobsoni* Oud. to acaricides', *Apidologie,* 30(2-3), 229-234.

Morgan, J. (2017) 'Operation Modomoc An experiment to test isolation status of the small Saltee island, Wexford',*The four seasons,* 67, 17-20.

McMullan, J. (2018) 'Adaptation in honey bee (*Apis mellifera*) Colonies exhibiting tolerance to *Varroa destructor* in Ireland', *Bee World,* 95(2), 39-43.

Momot, J. P. and Rothenbuhler, W. C. (1971) 'Behaviour genetics of nest cleaning in honeybees. VI. interactions of age and genotype of bees, and nectar flow', *Journal of Apicultural Research,* 10(1), 11-21.

Moretto, G., Goncalves, L. S. and Dejong, D. (1993) 'Heritability of Africanized and European Honey-bee defensive behaviour against the mite *Varroa jacobsoni*', *Revista Brasileira De Genetica,* 16(1), 71-77.

N.I.H.B.S. (2021) 'Statement about importing bees into Ireland' *The four seasons*, 81, 8-9

Nicodemo, D., De Jong, D., Couto, R. H. N. and Malheiros, B. (2013) 'Honey bee lines selected for high propolis production also have superior hygienic behavior and increased honey and pollen stores', *Genetics and Molecular Research,* 12(4), 6931-6938.

Oddie, M. A. Y., Neumann, P. and Dahle, B. (2019) 'Cell size and *Varroa destructor* mite infestations in susceptible and naturally-surviving honeybee (*Apis mellifera*) colonies', *Apidologie,* 50(1), 1-10.

Oleksa, A., Wilde, J., Tofilski, A. and Chybicki, I. J. (2013) 'Partial reproductive isolation between European subspecies of honey bees', *Apidologie,* 44(5), 611-619.

Page, R. E. and Fondrk, M. K. (1995) ' The effects of colony level selection on the social organization of honey bee (*Apis mellifera*) colonies – Colony level components of pollen hoarding', *Behavioral Ecology and Sociobiology,* 36(2), 135-144.

Page, R. E., Waddington, K. D., Hunt, G. J. and Fondrk, M. K. (1995) 'Genetic determinants of honey bee foraging behaviour', *Animal Behaviour,* 50, 1617-1625.

Pankiw, T. and Page Jr, R. E. (2000) 'Response thresholds to sucrose predict foraging division of labor in honeybees', *Behavioral Ecology and Sociobiology,* 47(4), 265-267.

Parker, R., Melathopoulos, A. P., White, R., Pernal, S. F., Guarna, M. M. and Foster, L. J. (2010) 'Ecological adaptation of diverse honey bee (*Apis mellifera*) populations (Adaptation in bees)', *PLoS ONE,* 5(6), e11096.

Pritchard, D., (2018) Geordie bees bite the mite, *S.I.C.A.M.M. conference abstract,* Finland.

Pritchard, J. K., Stephens, M. and Donnelly, P. (2000) 'Inference of population structure using multilocus genotype data', *Genetics,* 155(2), 945-959.

Rath, W. and Drescher, W. (1990) ' Response of *Apis cerana* Fabr towards brood infested with V*arroa jacobsoni* Oud and infestation rate of colonies in Thailand', *Apidologie,* 21(4), 311-321.

Rennie, J., White, P. B. and Harvey, E. J. (1921) '(1) The etiology of the disease', *Transactions of the Royal Society of Edinburgh,* 52(4), 737-754.

Rinderer, T. E., Harris, J. W., Hunt, G. J. and de Guzman, L. I. (2010) 'Breeding for resistance to *Varroa destructor* in North America', *Apidologie,* 41(3), 409-424.

Rinderer, T. E., Kuznetsov, V. N., Danka, R. G. and Delatte, G. T. (1997) 'An importation of potentially varroa-resistant honey bees from far-eastern Russia', *American Bee Journal,* 137(11), 787-789.

Rinkevich, F. D. (2020) 'Detection of amitraz resistance and reduced treatment efficacy in the varroa mite, *Varroa destructor*, within commercial beekeeping operations', *PLOS ONE,* 15(1), e0227264.

Rosenkranz, P., Aumeier, P. and Ziegelmann, B. (2010) 'Biology and control of *Varroa destructor*', *Journal of Invertebrate Pathology,* 103, Supplement, S96-S119.

Rowell, G. A., Taylor, O. R. and Long-Rowell, M. A. (1992) 'Spatial dynamics of the honey bee mating system (*Apis mellifera* L.)', *Journal of the Kansas Entomological Society,* 65(3), 218-222.

Scientific Beekeeping, (2018), Website: www.scientificbeekeeping.com accessed 17th Dec. 2018

Scott, D., (2014), Breeding for varroa resistance in Ireland, *The four seasons*, 53 (Spring 2014)

Seeley, T. D. (2007) 'Honey bees of the Arnot Forest: a population of feral colonies persisting with *Varroa destructor* in the northeastern United States', *Apidologie,* 38(1), 19-29.

Seeley, T. D. and Griffin, S. R. (2011) 'Small-cell comb does not control varroa mites in colonies of honeybees of European origin', *Apidologie,* 42(4), 526-532.

Simone-Finstrom, M., Walz, M. and Tarpy, D. R. (2016) 'Genetic diversity confers colony-level benefits due to individual immunity', *Biology Letters,* 12(3).

Smodiš Škerl, M. I., Kmecl, V. and Gregorc, A. (2010) 'Exposure to pesticides at sublethal level and their distribution within a honey bee (*Apis mellifera*) colony', *Bulletin of Environmental Contamination and Toxicology,* 85(2), 125-128.

Spivak, M. and Gilliam, M. (1998a) 'Hygienic behaviour of honey bees and its application for control of brood diseases and varroa - Part II. Studies on hygienic behaviour since the Rothenbuhler era', *Bee World,* 79(4), 169-186.

Spivak, M. and Gilliam, M. (1998b) 'Hygienic behaviour of honey bees and its application for control of brood diseases and varroa Part I. Hygienic behaviour and resistance to American foulbrood', *Bee World,* 79(3), 124-134.

Starr, C. (2006) ‘Steps toward a general theory of the colony cycle in social insects’, *Life cycles in social insects: Behaviour,Ecology and Evolution,* St. Petersburg: St. Petersburg University Press, pp.1-20.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) 'MEGA6: Molecular evolutionary genetics analysis version 6.0', *Molecular Biology and Evolution,* 30(12), 2725-2729.

Taylor, M. A., Goodwin, R. M., McBrydie, H. M. and Cox, H. M. (2008) 'The effect of honey bee worker brood cell size on *Varroa destructor* infestation and reproduction', *Journal of Apicultural Research,* 47(4), 239-242.

Tarpy, D. R. (2003) 'Genetic diversity within honeybee colonies prevents severe infections and promotes colony growth', *Proceedings of the Royal Society B-Biological Sciences,* 270(1510), 99-103.

Tarpy, D. R., vanEngelsdorp, D. and Pettis, J. S. (2013) 'Genetic diversity affects colony survivorship in commercial honey bee colonies', *Naturwissenschaften,* 100(8), 723-728.

Uzunov, A., Brascamp, E. W. and Büchler, R. (2017) 'The basic concept of honey bee breeding programs', *Bee World,* 94(3), 84-87.

Uzunov, A., Büchler, R. & Bienfeld, K., (2015) Performance testing protocol: A guide for European honey bee breeders, Vers. 1.0, SMARTBEES

Chapter 1 appendices BP I to BP VIII

APPENDIX BP I

Sugar Shaker test to determine *varroa* levels in a colony

List of equipment needed: 1 large clean dry bucket or basin, a second large, clean white bucket, a small amount of water will be placed in this bucket for each colony test hence the need for two buckets. A drinks bottle with clean water to refresh the white bucket after each mite count. 1 or 2 sugar shakers with mesh lid.

1. Pour about 1cm or half an inch of **clean water** into

the white bucket. If there is water from a test on another colony it should be poured out. **Ensure no mites are left from the previous test.**

2. Open the hive and pull out a frame from the **centre of the brood box or brood area**.

Ensure the queen is not on the frame, so she doesn't end up in the shaker. (Even if the queen ends up in the shaker by mistake, this method should not kill her but best to be safe).

3. **Shake the bees off the frame into the clean dry bucket**, keep tapping the bucket gently to keep the bees disorientated.

NB: If there are open honey cells, shaking the frame will release honey along with the bees and they will end up a sticky mess. Either choose an adjacent honey-free frame or **use a bee brush** and gently brush the bees into the dry bucket.

4. **Pour or scoop the bees into the shaker and close the mesh lid**. **Ensure the mesh lid is on securely.** Tap the shaker to ensure it is filled with bees up to the pre-marked 100 ml line.

5. If there are too many bees in the shaker, open the mesh lid slightly and let some fly off. Re-tap to check the level.

6. Add approximately **two tablespoons of icing sugar** to the shaker

6. Swirl the shaker around to ensure all the bees are well coated in the sugar. Placing the solid lid provided on top of the mesh lid prevents loss of sugar for easier coating of the bees.

7. **Leave for 5 minutes** to allow the mites to get off the bees. It is good tohave a second shaker so you can begin working on another hive.

8. Turn the shaker upside down over the white bucket and shake gently but firmly for one minute. Keep a finger on the edge of the lid to stop it falling off. Keeping the shaker below the level of the bucket rim will reduce the risk of mites blowing away in the wind.

9. The icing sugar will dissolve immediately and leave the mites floating, if mites are still falling towards the end of the first minute shake for another minute or until all mites stop falling.
Count the number of mites in the white bucket and record it on the inspection sheet. **For the percentage**, double the number of mites counted, and divide by 3.However if there is no brood present simply divide by 3

10. The bees can now be returned to the hive.

APPENDIX BP II – Inspection recording form

NUIG/NIHBS PROJECT TO BREED FOR VARROA TOLERANCE

Beekeeper name	
Date of inspection	
Apiary name	

If you are assessing more than one apiary please use a separate form for each one

Location Exact location not necessary	
Total colonies in apiary excl. nucs	

Please note: Only the strongest and weakest colonies along with the maximum of a further 8 that have been chosen randomly should be tested in each apiary.

Sample	Hive #	Number of mites found in 300 bees (100 ml)	% of mites*	Queen age in months	Queen from source External or Internal to apiary?	Date last treated	Treatment history over last 12 months
	eg: 4	eg: 6	eg: 4	eg: 15	eg: Internal	eg: Dec 2014	eg: Oxalic in Dec, Apiguard in Sept
Strongest							
Weakest							
1							
2							
3							
4							
5							
6							
7							
8							

	WORKER BROOD		DRONE BROOD		
Hive # as above	% of Sealed Brood Present On frame	% of Open Brood Present On frame	% of Sealed Brood Present On frame	% of Open Brood Present On frame	Please detail if anything happnened since last treated
	eg; 65%	eg: 10%	eg: 15%	eg: 10%	eg: Hive was split in May but queen always laying, no swarm

*Divide No. of mites found by 3 and multiply by 2 e.g. (6 mites found/ 3) x 2 = 4% mite infestation.If no brood do not mulitply by 2

Feb-16

	Colony characteristics based on usual or average behaviour		**Rating system**			***Propolis rating**
			0 = Unsatisfactory			0 = High
			1 = Satisfactory			1 = Med/High
			2 = Good			2 = Medium
			3 = Very good			3 = Med/Low
			4 = Excellent			4 = Low
			5 = Outstanding			5 = Very low/None
Hive # as overleaf	**Docility**	**Steadiness**	**Brood pattern**	**Pollen storage**	**Comb building**	***Amount of Propolis**
	eg: 4	eg: 5	eg: 4	eg: 3	eg: 4	eg: 4

APPENDIX BP III

INSPECTION FORM COMPLETION NOTES

HIVE CHOICE

Previously submitted varroa counts – Please use the same colonies as before.

Apiary with less than 10 hives – Inspect all hives

Apiary more than 10 hives – Choose your best and worst hives. Choose another 8 hives randomly.

LOCATION

The exact hive location is not essential however please provide the townland if you can.

If the apiary is moved during the year, please provide the **location where it is most of the time or during the greatest period of nectar flow**.

APIARY NAME

Please use a separate form for each apiary if you are testing more than one apiary.

TOTAL NUMBER OF COLONIES

The total size of the apiary at the time of testing

CALCULATION OF MITE INFESTATION

We only require the **number of mites you count in 300 bees** and whether or not brood is present.

Brood present on frame : Number of mites divided by 3 **x 2** = **Percentage of mite infestation**
E.g. 3 mites ÷ 3 **x 2**= 2.00 % mite infestation

No brood present on frame do not multiply by 2.
E.g. 6 mites ÷ 3 = 2.00 % mite infestation

If the percentage is 2% or less do not treat for varroa if you are happy to do so.
If the percentage exceeds 2% treat for varroa as you normally would. Re-queen from low varroa stock.

QUEEN AGE

Please enter this to the nearest month

WHERE QUEEN WAS SOURCED FROM

Internal – Colony queened from within the current apiary
External – Colony queened from another apiary, even the other apiary belongs to you

BROOD VALUES - SAMPLE FRAME ONLY

The percentage of brood should be given as a percentage of the entire area of both sides of only the frame from which the 300 bees were taken. If you need two frames to get enough bees use the one from which you got most bees for the calculation.

Brood percentage is subjective and can be difficult to estimate therefore it is best for you to have a standard guide from which to compare each frame and always use that method.

For example, Use the base of the sugar shaker tub (870 ml) which has an area of 64 sq cms.
The base of the tub is 1/11 or 9% of the National frame and 1/14 or 7% of the Commercial frame.

National frames			Commercial frames	
Brood covers shaker tub bases	% of brood frame		Brood covers shaker tub bases	% of brood frame
1	9%		1	7%
2	18%		2	14%
2.5	22.5%		2.5	17.5%
3	27%		3	21%
3.5	31.5%		3.5	24.5%

Once you have estimated the brood percentage please give your best estimation of the ratio of Open brood to Closed brood. (eg: 3 to 1; 1 to 2) Always give the Open brood figure first. Finally tick one box to indicate, in your opinion, the level of combined open and closed drone brood on the frame.

Any queries please contact Keith Browne at nuigbeeresearch@gmail.com, 091 494490 or 089 2004762

APPENDIX BP IV – Inspection method

Dear beekeeper,

Welcome to the 2016 and the second full year of the joint NIHBS/NUIG breeding programme to increase the prevalence of honey bee colonies in Ireland that are tolerant of the parasitic mite, *Varroa destructor*.

The programme continues into another year with the support of all parties involved because it is seen as part of the best long-term approach to assist *Apis mellifera mellifera* in a recovery from *varroa*. NUIG look forward to another year assisting beekeepers in the first moves towards improving the *varroa* tolerance of their honey bee population however additional participants are essential if the programme is to yield a successful outcome for NIHBS members and other beekeepers.

A. EXISTING BEEKEEPERS: DATA COLLECTION FOR 2016

If you have **already submitted** *varroa* counts, thank you. Please **continue to do so using the same colonies** as before. We ask that all beekeepers who submit counts also include the very important **additional information** about the colonies. These are necessary to understand their characteristics in order that desired traits can be kept in the population.

VERY IMPORTANT; PLEASE NOTE : This year the *varroa* counts will take place in two discrete windows of time as leaving it open-ended last year appeared to create some confusion.

COUNT 1: Between the **7th and 22nd of May but before any anti-varroa treatment.**
If you treat early, conduct the count before treatment irrespective of the date.

COUNT 2: Between the **13th and 28th of August but before anti-varroa treatment.**
If you treat early, conduct the count before treatment irrespective of the date.

Please conduct your count, complete the additional information in the inspection form, which has been revised to further simplify the process and submit it by either;

email to nuigbeeresearch@gmail.com (There is an Excel version available for this)

Or

post to Keith Browne, Room 203 Ryan Institute, Department of Zoology, NUI Galway, University Road, Galway.

BEE SAMPLES

If you have **already sent samples of bees** to either NUIG, University of Limerick or Limerick Institute of Technology there is no need to send any further samples for the moment.

If you have not submitted samples or this is your first year taking part we would appreciate a sample of **10 bees from each colony** from which *varroa* counts were done. An empty matchbox should suffice as a container and protect the bees in transit.

The humane euthanasia of bees is to place them in a fridge for a short while before freezing them. Some form of insulation also helps to keep them fresh whilst they are in transit.

Each individual sample of bees should be marked with the following four pieces of information please:

Date of sampling	**Apiary name**
Beekeeper name	**Hive number**

Finally, make sure your **contact address and/or phone** number is included in the package.

B. INTERESTED IN GETTING INVOLVED?

If a significant number of beekeepers in Ireland conduct selective breeding for *varroa*-tolerance then it is possible that the result will be an increase in the tolerance to *varroa* across the entire honey bee population. Not all colonies would have to be *varroa*-tolerant for a population-wide effect.

If you keep *Apis mellifera mellifera*, **even if you think they may be hybridised** to some extent, your involvement is required.

The procedure is quite straightforward and is described in detail in specific documents. Here are the main points:

1. At the same times **twice a year assess the percentage of *varroa*** in your colonies using the sugar shaker method as described.
2. If the colony's percentage is **2% or under, preferentially use it to breed** from. Additionally, **don't treat** for *varroa,* provided you are happy not to, in order to determine if the low percentage results from a colony trait rather than anti-*varroa* treatment.
3. If the percentage is **over 2%** you **can treat** for *varroa* as you normally would and preferably **not breed** from this colony or colonies. **Try re-queen** from low *varroa* stock

We are looking for data from a **maximum of 10 colonies per apiary** however there is nothing to stop you from applying this protocol to your entire apiary. We are **also interested in data from apiaries under 10 colonies**.

Even if you decide recording the information and returning it to us is not possible for you, you can still use this selective breeding method and hopefully help increase the level of *varroa* tolerance in the Irish honey bee population.

If you are interested in taking part in the *varroa* counts and data collection please read the additional attachments carefully, particularly the inspection form and sugar shaker method. If you then decide you **understand the process,** can **spare the time** and wi**sh to proceed, please contact us as there are a limited number of places** in the experimental core group due to genotyping costs.

That said, if you wish to conduct the counts anyway and submit them we would be pleased to have the data, however we cannot guarantee your bees will be genotyped.

To conduct counts, **obtain one or two sugar shakers** from either your usual beekeeping equipment supplier or by contacting NIHBS and then follow the instructions in **A** above for existing beekeepers.

APPENDIX BP V

Aims, objectives and overview of the breeding programme and accompanying research

Primary aim*: To reduce the dependency on chemical treatment for varroa control by increasing the prevalence of varroa tolerant colonies in Ireland.*

To augment the feral/wild native Irish honey bee population with varroa tolerant colonies.

The programme aims to reduce the amount of chemicals used to treat honey bees for varroa infestation by increasing the number of colonies that can tolerate varroa without beekeeper intervention. It is hoped that this can be achieved using selective breeding in much the same way beekeepers have done for traits such as docility and honey production.

The chief objective is to assist beekeepers to use a method by which they can achieve this. The more beekeepers that use the breeding method, the better. Within this, a core group will submit detailed reports on their colonies to NUIG to help the research into varroa tolerance in Irish honey bees. These core group bees will be genotyped in order to obtain a genetic picture of the Irish honey bee population.

Simply put, colonies are assessed for varroa and selection for breeding uses those assessments. Those with low varroa levels are not treated for the mite and are preferentially used to breed from. Those colonies with varroa levels above a certain threshold are treated and, when appropriate, re-queened with the offspring of low-varroa queens from within the same apiary or from other participating beekeepers. This is not a fool-proof method, there is no such thing, therefore it is important that some free, altruistic, re-queening forms part of this programme in order to help beekeepers who lose colonies by not treating. In other words, beekeepers need to assist one another if the breeding programme is to help everyone.

Meanwhile research at NUIG will endeavour to uncover how some honey bees in Ireland can tolerate and even resist varroa mites. This information will be fed back into the breeding programme to help steer the selection process. The research and the breeding programme also aim to assist the native Irish honey bee's feral and wild population by augmenting it with colonies bred from varroa tolerant queens.

APPENDIX BP VI - FAQs for NIHBS/NUIG varroa Tolerance Breeding Programme

Q. Who do I contact in NUIG?
A. Keith Browne. Tel. 091 494490/ 089 200 4762 or nuigbeeresearch@gmail.com

Q. Who do I contact in NIHBS?
A. Your regional organiser: leinster@nihbs.org, ulster@nihbs.org, munster@nihbs.org, Connaught@nihbs.org

Q. How often must I do the varroa count?
A. Twice a year.

Q. When must I do the count?
A. Before you treat for varroa.
Ideally, Count 1 between 7th & 22nd May & Count 2 between13th & 28th August.

Q. Where can I learn how to do the Sugar Shaker method?
A. Keith Browne will demonstrate Sugar Shaker method & give talk on programme at OBKA apiary in Charleville estate, Tullamore on 15th May at 1.00 p.m. All welcome.

Q. How long does it take to do the Sugar Shaker test?
A. A few minutes per hive. It could add about 10 minutes to your inspection of a hive.

Q. Do I have to send bees away?
A. Yes, a sample of 10 bees per hive to be sent to NUIG in a matchbox.

Q. Do I send live bees?
A. No, please send dead bees.

Q. How do I prepare bees for posting?
A. Label matchbox with apiary name, hive no. beekeeper name and date of sampling.
Collect bees. On your return home, put matchbox in fridge for approx. 20 minutes then into freezer. Use padded envelope to send full matchboxes to NUIG. Be sure that your contact address and/or phone number is included in the package.

Q. Why can't I put bees straight into freezer?
A. It is more humane to use fridge first. Also, if bees go straight into freezer, there will be a blob of frozen/defrosted honey in matchbox. Avoid this by using fridge first.

Q. How do I label matchbox?
A. Use biro or permanent marker. Writing should be clearly legible after freezing.

Q. Will training be offered around the country?
A. NIHBS events and workshops will usually offer training. Please request your Regional Organiser.

Q. Where do I get tubs for the Sugar Shaker test?
A. Your Regional Organiser should supply them. If not, contact NIHBS secretary nihbs.secretary@gmail.com

Q. Will my bees be genotyped?
A. Only bee samples submitted by the core group of beekeepers can be genotyped due to the costs involved.

Q. How do I assess my colony's characteristics?
A. The assessment is based on the Hive record form developed by Galtee bee breeding group. It should reflect the usual characteristics of the colony, not just what you see on the day of assessment. The list of what to look for is on the inspections form and shown below.

DOCILITY: Non-jumping, non-stinging, non-following.
STEADINESS: Absence of running on comb
BROOD PATTERN: Brood compactness, Absence of empty cells.
POLLEN STORAGE: Pollen packed over, around and under broodness.
COMB BUILDING: Speed in occupying supers, drawing foundation, honey and quality of comb capping.
These characteristics are rated 0 to 5, from unsatisfactory to outstanding, as shown on the inspection form.

APPENDIX BP VII

Observation metrics for sequence chromatogram rating

The observations are made at the default height and width settings under which the chromatogram is first opened. Alteration to the settings can be made later for accurate determination of the clarity of individual peaks.

Trace clarity

5: Clear medium to strong signal over entire length, no background signal interference.
4: Clear low to medium signal over>75% length, little or no background signal interference.
3: Slightly noisy signal under 50% of length with low background signal interference.
2: Noisy signal over 50% of length with low/medium background signal interference.
1: Noisy signal over 50% of length with medium/strong background interference.
0: Very noisy signal over 75% of length with strong background interference. Illegible by eye.

Signal strength

Strong:	Peaks over half height
Medium:	Peaks roughly to half height
Low:	Peaks below half height
Very low:	Peaks discernible but just above baseline
Extremely low:	No peaks discernible. System still provides a character output
None:	No peaks and no character output.

APPENDIX BP VIII: Mite percentages recorded for all colonies for which data was received from 2013 to 2016b along with presence/absence data on treatment for the 12 months prior. (Yes/No/Not recorded)

Beek	Colony	2013	Treat	2014	Treat	2015.00	Treat	2016a	Treat	2016b	Treat
1	B1					2.00	y	4.00	y		
1	B2			4.00	nr	14.67	y	1.33	y		
1	B3			2.67	nr	2.00	y	2.67	y		
1	B4			2.00	nr						
1	B5			1.33	nr	0.00	y	3.33	y		
1	B6			0.00	nr	2.00	y	0.00	y		
1	B7							0.67	y		
1	K1			2.00	nr	4.00	y	4.00	y		
1	K2			2.67	nr	3.33	y	4.67	y		
1	K3					4.00	y				
1	K4			6.67	nr	2.67	y	0.67	y		
1	K5			2.00	nr	1.33	y	2.67	y		
1	K6			1.33	nr	7.33	y				
2	94	0.00	y	12.00	y	4.00	y	0.67	y		
2	100	1.70	y								
2	104	1.30	y								
2	111	0.30	y								
2	105	1.00	y								
2	110	4.30	y	6.67	y						
2	119			12.67	y	42.00	y				
2	141			46.67	y	3.33	y	4.00	y		
2	160			33.33	y						
2	161			2.00	y			0.00	y		
2	163			1.33	y			0.00	y		
2	166			10.00	y						
2	168			16.00	y	4.00	y				
2	AH1			12.00	y						
2	BC1	3.00	y	2.67	y			2.67	y		
2	DA7	0.70									
2	DA8	0.00	y	20.00	y	17.33	y	0.00	y		
2	G4			52.00	y	2.67	y	0.00	y		
2	PD1	0.30	y	26.67	y	20.67	y				
2	240-15							3.33	y		
2	212-15							0.00	y		
2	G66							0.00	y		
2	211-15							0.00	y		
2	190							0.00	y		
2	205-15							0.67	y		
2	JBG3							0.00	y		
2	JBG8							0.67	y		
2	JBG1							0.00	y		
2	JBG2							0.00	y		
2	JBG4							0.67	y		

2	JBG5							6.67	y		
2	JBG6							0.00	y		
2	JBG7							1.33	y		
3	A1									0.67	y
3	C1					15.33	y			0.67	y
3	C2					0.00	y			3.33	y
3	C3					2.00	y			0.00	y
3	C4					6.00	y			0.00	y
3	C5					4.00	y				
3	C6					0.00	y			0.00	y
3	C6A					0.67	y				
3	C7					0.00	y				
3	C8					0.67	y				
3	Pt1									1.33	y
3	Pt2									0.00	y
3	Pig1					8.00	y				
3	Pig2									0.67	y
3	Pig3									0.00	y
4	D1					10.00					
4	D2					16.67					
4	F17					12.00					
5	GordN9					0.00	y				
5	GordN10					0.00	y				
5	GordN11					0.67	y				
5	GlenL5			1.33	y	0.00	y				
5	GlenL6			0.00	y						
5	GlenL7			0.00	y	0.00	y				
5	GlenL9			0.00	y						
5	GlenL10			0.00	y	0.00	y				
5	GlenL11			0.67	y						
5	GlenL12			0.00	y	0.67	y				
6	1	1.00	y	0.00	y	1.33	y	0.67	y	1.33	y
6	2	2.00	y	4.00	y	1.33	y			0.00	y
6	3	0.00	y	0.67	y			1.33	y	0.00	y
6	4	0.00	y	2.67	y	4.00	y			0.00	y
6	5	1.00	y	2.67	y					1.33	y
6	6	1.00	y	0.67	y	1.33	y			0.67	y
6	7	1.00	y	8.67	y						
6	8	3.00	y	9.33	y						
6	9	1.00	y	5.33	y	0.00	y				
6	10	14.00	y	2.00	y						
6	11	1.00	y			0.67	y				
6	12					2.00	y				
6	13			0.00	y						
6	14			0.00	y						
6	15			0.67	y	0.00	y				

7	BA1						0.00	y	20.00	y
7	BA10						6.00	n	0.00	y
7	BA2						0.00	y	8.00	y
7	BA4						4.00	y	30.00	y
7	BA7						1.33	y	10.00	y
7	BA9						0.67	n	36.00	y
7	DA1		2.00	y	2.30	y			0.00	y
7	DA2		3.33	y	4.30	y			0.00	y
7	DA3		6.00	y						
7	DA4								0.67	y
7	DA5		12.67	y					0.67	y
7	DA6		2.00	y					2.00	y
7	DA7		2.67	y						
7	DA8				20.00	y				
7	DA9		2.00	y						
7	DA10				2.30	y				
7	DA11		12.00	y						
7	DA12		6.00	y						
7	Glen3		14.00	y						
7	DR1				17.00	y				
7	DR2				4.00	y				
7	DR3				3.00	y				
7	DR4				4.00	y	1.33	n	2.00	y
7	DR5				3.00	y	0.67	n	4.00	y
7	DR8				5.00	y				
7	DR7								0.67	y
7	DR11								0.00	y
7	DR17								1.33	y
7	GR1						1.33	n	6.67	y
7	GR4						1.33	n	5.33	y
7	GR6						0.67	n	12.67	y
7	GR7						13.33	n	10.67	y
7	GR8						0.67	n	4.67	y
7	NU1								4.00	n
7	NU2A								8.00	n
7	NU5						3.33	n	10.00	y
7	NU6						4.67	n	5.33	y
7	GR5						4.00	n		
7	GR2						0.00	n		
7	GR10						2.00	n		
9	9		3.33							
9	5		40.00							
9	11		18.00							
10	1				0.67	y				
10	26				0.00	y				

10	27			0.00	y				
10	2			2.00	y				
10	28			18.66	y				
10	3			6.00	y				
10	29			2.00	y				
11	1					0.00	y	4.67	y
11	7					0.00	y	8.00	y
11	13					0.00	y	8.67	y
11	34					0.00	y	5.33	y
11	24					0.00	y	4.00	y
11	111					0.00	y	4.67	y
11	5					1.33	y	7.33	y
11	15					0.00	y	4.67	y
12	1			0.67	y			1.33	y
12	2			0.67	y			2.67	y
12	5			4.00	y			1.33	y
12	7			3.33	y			1.33	y
12	8							8.00	y
12	12			0.67	y			0.00	y
12	14			4.00	y			1.33	y
12	16			2.00	y			0.00	y
12	20			2.67	y				
19	7			4.00	y				
19	5			13.33	y				
19	11			2.67	n				
19	10			0.67	n				
19	9			5.33	y				
19	6			6.00	n				
19	5			0.67	n				
19	4			8.67	n				
19	2			4.67	n				
19	1			2.00	n				
26	LG1					0.00	y	5.33	y
26	LG2					2.67	y	4.67	y
26	LG3					24.00	y	0.00	y
26	LG4					10.00	y	12.00	n
26	LG5					0.00	y	0.00	y
26	D1					0.00	y	3.33	y
26	D2					0.67	y	6.00	y
26	D3					0.00	y	0.00	y
26	D4					0.00	y	2.00	y
26	D5					0.00	y	7.33	y
27	5A			1.33					
27	5B			0.33					

27	5				5.33			
27	22				2.00			
27	21				2.67			
27	2				4.00			
27	Dan7				1.00			
27	19				4.00			
27	23				0.00			
27	11		53.33	Nr	0.00			
27	16		5.33	y	0.00			
27	9		28.67	y	2.67			
27	1		0.67	y				
27	18		2.00	y	4.67			
27	6		2.67	y	40.67			
27	8		1.33	y				
27	7		1.33	y				
27	13		2.00	n				
27	5		1.33	n				
27	20		2.00	n				
27	19		8.00	n				
27	3		2.67	y				
27	4		2.67	y				
27	15		8.00	y				
27	12		5.33	n				
27	10		1.33	n				
27	2		1.33	n				
27	17		26.67	y				
27	14		8.00	y				
27	21		10.67	y				
28	91		1.33	y	2.67	y		
28	224		0.67	y	2.00	y		
28	102				0.00	y		
28	NAT		0.00	y	0.67	y		
28	42				0.00	y		
28	217				0.67	y		
28	215		0.00	y	0.67	y		
28	38				0.00	y		
28	220				0.67	y		
28	53				0.67	y		
28	1A				0.00	y		
28	85				0.00	y		
28	17		1.33	y	13.33	y		
28	11		0.67	y	19.33	y		
28	47				7.33	y		
28	3202		0.00	y	3.33	y		

28	223			0.00	y				
28	31			0.00	y				
29	1					2.67	y		
29	2					5.33	y		
29	3					3.33	y		
29	4					6.67	y		
30	A1					1.33	y		
30	C1					14.00	y		
31	6							0.00	y
31	5							0.00	y
31	1							0.00	y
31	2							0.00	y
31	3								
31	7							3.33	y
31	8							0.00	y
31	9							0.00	y
31	10							0.00	y
32	1							2.00	y
32	5							0.00	y
32	2							0.00	y
32	3							3.33	y
32	4							1.33	y
33	752 (864)							0.67	y
33	696 (842)							0.00	y
33	693							1.33	y
33	700 (872)							0.67	y
33	771							0.00	y
33	751							0.00	y
33	778							0.67	y
33	779 (863)							0.00	y
33	695							0.00	y
33	548 (870)							0.67	y
33	667							0.67	y
33	807 (855)							0.00	y
33	669(871)								
33	603								
33	661(862)								
33	666								
33	633								
34	P1							2.67	y

34	H5				6.00	y		
34	P2				3.33	y		
35	1				0.00	y		
35	2				0.67	y		
35	3				0.00	y		
35	4				0.67	y		
35	5				0.67	y		
35	6				0.00	y		
35	7				0.00	y		
35	8				1.33	y		
36	5				0.67	y	2.67	y
36	1				1.33	y		
36	4						1.33	y
36	2				0.00	y	19.33	y
37	1				0.67	y	4.00	y
37	2				0.00	y	5.33	y
37	3				0.00	y	8.00	y
37	4				2.67	y	4.67	y
37	5				0.00	y	9.33	y
38	F1				0.00	y	0.00	y
38	F2				3.33	y	1.33	y
38	F3				0.00	y	0.00	y
38	CDB1				1.33	y	2.00	y
38	L2				0.00	y	0.00	y
39	R4				0.00	n		
39	R2				9.33	y		
39	T5				0.67	n		
40	H1A				1.33	y	3.33	y
40	H3B				0.00	y	1.33	y
40	H2A				0.00	y	7.33	y
40	H3A				0.00	y	0.67	y
40	H7A				2.00	y	16.00	y
41	H4				2.00	y		
41	H7				0.00	y		
41	H2				3.33	y		
41	H9				0.00	y		
41	H11				0.00	y		
41	H5				2.00	y		
42	13				8.00	y		
42	14				2.00	y	30.00	y
42	5				0.67	y	0.67	y
42	22				1.33	y	13.33	y
42	6				0.00	y	0.00	y
42	25				0.67	y		
42	36				0.00	nr	2.00	y
43	2				8.00	y		

43	1				3.33	y		
44	1				0.00	y	10.67	y
44	4				0.00	y	6.67	y
44	3				0.67	y	4.67	y
44	2				0.67	y	3.33	y
45	2				2.00	y		
45	1				3.33	y		
45	3				3.33	y		
45	4				2.67	y		
46	HE				4.00	y		
46	HB				16.67	y		
46	H27				8.00	y		
46	H15				2.00	y		
46	H1				20.67	y		
46	H4				8.00	y		
47	1				0.00	y	0.67	y
47	2						6.67	nr
47	3						4.67	nr
49	T4				2.67	y	8.00	y
49	T6				0.67	y	2.00	y
49	T1				0.00	y	0.00	y
49	T2				0.00	y		
49	T3						0.67	y
50	2				0.00	y		
50	1				0.67	y		
50	5				0.00	y		
50	3				0.00	y		
51	4				0.00	y		
51	3				0.00	y		
51	2				0.00	y		
52	43				2.00	y		
52	60				5.33	y		
52	38				5.33	y		
52	35				0.00	y		
52	90				1.33	y		
52	23				0.67	y		
52	68				1.33	y		
52	9				0.00	y		
52	37				0.00	y		
52	66				0.00	y		
53	4				2.00	y		
53	3				2.00	y		
54	D1				0.00	y	0.00	y
54	D2				0.00	y		
54	D3						1.33	y
54	K1				0.00	nr	0.00	y

54	K2				0.00	nr	1.33	y
55	B1				2.67	y		
55	H1				3.33	y	3.33	y
55	R1				3.33	y	6.67	y
55	T1				4.00	y	8.67	y
56	1				0.00	y	0.00	y
56	6				0.00	y		
56	3							
56	5				0.00	y	8.00	y
56	11				0.00	y	0.00	y
56	8				0.00	y	16.67	y
56	10				0.00	y	3.33	y
57	H1				2.67	y	11.33	y
57	H2				11.33	y	16.00	y
57	H3				15.33	y	18.00	y
57	4				4.67	y	18.67	y
57	5				0.67	y		
57	9				0.00	y	14.67	y
58	8				0.00	y		
58	4				0.67	y		
58	1				0.00	y		
58	3				0.67	y		
58	5				0.00	y		
58	6				0.00	y		
58	7				0.67	y		
59	4				4.67	y	5.33	y
59	6				0.00	y	1.33	y
59	2				1.33	y	6.67	y
59	5				2.00	y	2.67	y
59	1				2.67	y	1.67	y
59	3				4.67	y	2.00	y
60	2				1.33	y		
60	5				0.67	y		
60	1				0.00	y		
60	3				0.00	y		
60	4				0.00	y		
60	6				0.67	y		
61	K2		0.00					
61	K3		4.00		1.33	y		
61	K4		0.67		2.00	y		
61	K5		5.33					
61	K6		12.67		0.67	y		
61	K7		1.33		0.67	y		
61	K8		2.00		1.33	y		
61	K9		14.67					

61	K10		5.33		0.67	y		
61	K11		2.00					
61	K12		15.33					
62	3						6.67	n
62	1						0.67	n
62	6						2.00	n
62	2						2.67	n
62	4						13.33	n
63	2KL				1.33	y	2.00	y
63	1KL						0.00	nr
63	4KL						1.33	nr
63	3KL				0.00	y	6.00	y
65	3						0.00	
65	6						10.00	y
65	1						3.33	y
65	2						9.33	y
65	5						8.67	y
66	RG1						3.33	n
66	RG2						9.33	y
66	RG3						1.33	n
Totals	**419**	**21**	**98**	**116**	**226**		**172**	

Chapter 2

Determining the status of free-living honey bee,

Apis mellifera L, colonies in Ireland.

This section was revised and published as:

Keith A. Browne, Jack Hassett, Michael Geary, Elizabeth Moore, Dora Henriques, Gabriele Soland-Reckeweg, Roberto Ferrari , Eoin Mac Loughlin , Elizabeth O'Brien, Saoirse O'Driscoll, Philip Young, M. Alice Pinto & Grace P McCormack (2020): Investigation of free-living honey bee colonies in Ireland, *Journal of Apicultural Research*, DOI:10.1080/00218839.2020.1837530

My involvement included database design and management, sample collection, DNA preparation and PCR, data analysis, production of figures, tables and networks and writing the majority of the manuscript

ABSTRACT

Apis mellifera mellifera, the dark northern honey bee, is generally considered to be extinct in the wild over most of its range because of replacement by and hybridisation with other sub-species, the impact of the ectoparasitic mite *Varroa destructor* and the effects of habitat loss and agricultural pesticides. In Ireland, the furthest west of the subspecies range, a managed population of pure *Apis mellifera mellifera* was recently shown to exist however the unmanaged population had yet to be investigated. It is shown here that unmanaged free-living honey bee colonies are present and widespread in Ireland, inhabiting a mixture of nesting habitats. The potential of individual colonies persisting naturally and unaided over multiple years is investigated. The population is further described using mitochondrial, microsatellite and SNPs evidence. Finally, the implications of conserving this population and its possible role in improving the fitness of the managed population both in Ireland and the rest of its European range is discussed.

INTRODUCTION

Varroosis caused by the invasive ectoparasitic mite *Varroa destructor* (Anderson and Trueman 2000) (varroa), combined with increasing introgression and replacement with introduced sub-species and hybrid strains, changes in land use and the proliferation of pesticide use has left some sub-species of *Apis mellifera* in a state of near extinction across much of their range (De la Rua *et al.* 2009). In northern Europe, from Ireland through France and Germany, eastwards to the Ural Mountains of Russia, northwards into Scandinavia and southwards to the Alps, M-lineage *A. m. mellifera* is the native sub-species (Ruttner 1988). It is locally known alternately as the dark Northern honey bee, brown bee or black bee. C-lineage sub-species such as *A. m. carnica* and *A. m. ligustica* and cross-lineage hybrids such as Buckfast, introduced into the managed M-lineage population have altered its genetic integrity (Jensen *et al.* 2005, Soland-Reckeweg *et al.* 2009, Pinto *et al.* 2014, Parejo *et al.* 2018, Ellis *et al.* 2018) and left the strong possibility that genes for locally adapted traits may have been removed from the population (Randi 2008). Even where *A. m. mellifera* has been retained, such as in pockets in France or Læsø island in Denmark, there is increased probability of reduced genetic diversity as a consequence of beekeepers' legitimate efforts to keep colony numbers up by raising multiple queens from single colonies and conducting colony splits (De la Rua *et al.* 2009). In the midst of efforts to address issues there has been insufficient investigation into the status of the free-living population (Moritz *et al.* 2005), leaving considerable uncertainty about its current state and conservation need. As free-living honey bees represent a significant proportion of the population that is adapting to selective pressures placed on them, particularly varroa infestation, it seems wise to characterise such bees where they persist, since these may well form the basis of future managed colonies.

More than 95% of beekeepers in the Republic of Ireland consider themselves as hobbyists (Chauzat *et al.* 2013) with the result that there is little official interest in widescale monitoring of the honey bee population. Discussions with conservation and beekeeper groups indicated that they understood that any free-living colonies that did exist must be hybridised with C-lineage sub-species and hybrids such as *A. m. ligustica* and Buckfast respectively, the same as managed colonies. Furthermore, they felt that the free-living colonies did not survive more than one or two years and therefore the picture was one of constant re-colonisation of cavities by swarms from the surrounding managed population rather than a perpetuation of

queen lineages through supersedure or swarming. This in turn, indicated that the free-living population was feral and effectively a by-product of beekeeping. In the UK, the nearest country to Ireland in which *A. m. mellifera* also is native, both managed and feral bees were found to be heavily introgressed with C-lineage genotypes, although there was a small but significant difference between the two populations and feral lineages did not survive beyond 2.5 years (Thompson 2012). The convention of referring to honey bee colonies that do not reside within an apiary as feral has strong implications for the manner by which the public, beekeepers, government agencies and conservation bodies view them. It can influence their subsequent treatment by these groups since they are perceived as simply escapees from managed apiaries. There is no evidence to suggest that all are a result of escapes from managed hives nor, conversely, that any of these colonies were always wild. It is likely that both elements of the larger population mate freely with each other. As we cannot yet determine if the colonies we refer to here are wild or feral, we propose the use of the term "free-living" for colonies not managed by beekeepers in any form, irrespective of the habitat or type of site chosen for colony settlement.

There is a small but increasing number of beekeepers who are reporting non-treatment for varroa on managed colonies and sustaining losses similar to or even lower than among their treated colonies (Swindon honeybee conservation group 2017, Pritchard 2018, McMullan 2018). While these reports may warrant further investigation, it appears that there is a small but significant groundswell movement within the beekeeping community to breed from colonies that indicate some form of natural resistance to varroa. Any free-living colonies that can be shown to have exhibited multi-year survival without chemical treatment beyond the year 3 peak post-infestation losses observed from field studies (Korpela *et al*. 1992, Fries *et al*. 2006) are also likely to create considerable interest for beekeepers both in Ireland and the rest of Europe, particularly if they are pure *A. m. mellifera*. The possibility remains that genotypes not found in managed hives, possibly even representing locally adapted ecotypes, may yet be found in free-living colonies.

In Europe, free-living honey bees are generally understood to be extinct or near extinct in all but a handful of conservation areas and nature reserves such as the Hainich forest and Swabian Alb Biosphere reserve in Germany (Moritz *et al.* 2007, Kohl and Rutschmann 2018),

although the extent of research into their existence is limited. In Ireland, where research into the status of the black honey bee has been limited, the idea of their extinction is primarily based on the anecdotal evidence that wild honey bee colonies became absent from places where they used to be common, and few had been seen in recent times. In contrast to this speculation, there were indications, again anecdotal, that putative free-living *A. m. mellifera* colonies may exist in Ireland, including a small number of personal reports from beekeepers and members of the public to GPM and KAB. This added weight to genotyping evidence which indicated that there were more colonies present in an area in Ireland than could be accounted for by managed colonies only (Jaffe *et al.* 2010). For all other countries sampled (apart from one location in Italy), the numbers of colonies estimated from the genotyping approach matched with the number of managed colonies documented, indicating a loss of wild colonies at these locations (Jaffe *et al.* 2010)

We considered that this study was essential to inform stakeholders in honey bee conservation both in Ireland and worldwide. There was an imperative to determine if the purity of the Irish *A. m. mellifera* population as a whole (Hassett *et al.* 2018, Henriques *et al.* 2018) was reflected in the free-living population and also to determine their status in terms of distribution, survival and genetic diversity. Further impetus is added if one includes the possibility that the future of beekeeping in Ireland may ultimately rely on colonies that can survive varroa without any management, the likes of which may lie unrecorded in free-living colonies.

MATERIALS AND METHODS

Locating free-living colonies

In August 2016 a nationwide appeal for information on the location of free-living honey bee colonies was made via a press release through The Irish Times, a national newspaper. A request was also made on social media accounts, which had been initiated one year previously to allow an accumulation of followers, including many beekeepers. The story was also picked up by national television, which helped to strengthen the appeal. Reports were gathered by telephone, email and social media contact. Reports determined to be honey bee swarm capture or other bee species were excluded. Ambiguous initial reports were contacted to ensure all positive sightings were correct and added to earlier reports from Autumn 2015.

Colony sampling and monitoring

Samples of honey bees were collected from entrances of 76 free-living colonies across Ireland (Figure 1) using a combination of long-handled butterfly nets, clear glass jars and a proprietary "Bug buster" suction tube. Some colonies were sampled by the member of the public who reported the colony (the custodian). They were given instructions for sample collection and preservation prior to shipping to us. Samples were cooled immediately upon capture before storage at -20^{o}C until DNA extraction. From the sampled colonies, those where continuing contact could be made with the custodian were monitored by means of a short telephone questionnaire in the Spring and late Autumn of each year to determine the level of colony activity, swarming activity and a cause of death if applicable. Custodians also had contact details for us to report any activity they felt was of interest. From nests where activity indicated colony survival, further bees were obtained where possible.

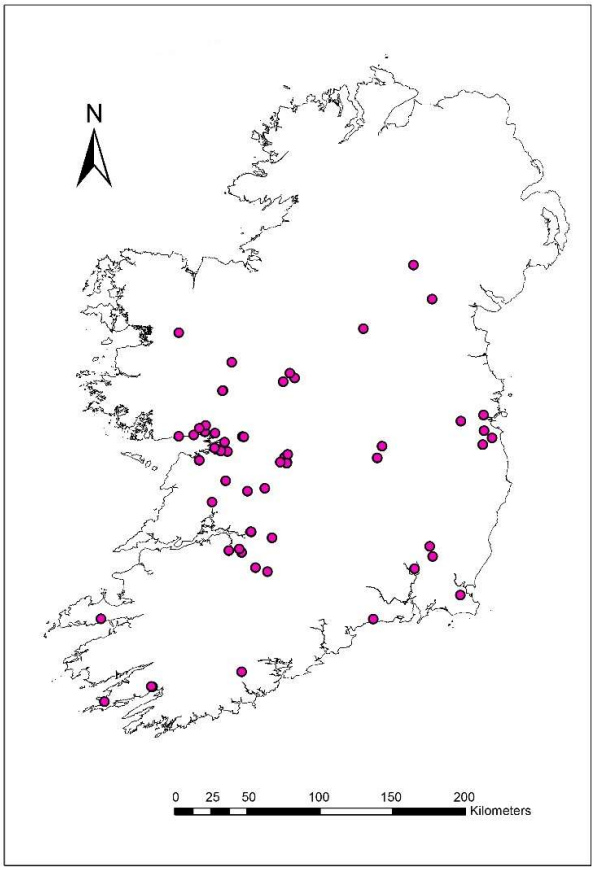


Figure 1. Locations of the 76 free-living colonies sampled in Ireland between 2015 and 2018 from the 182 reported. Some locations contained multiple colonies.

Varroa and Nosema testing

On the first year of sampling (2016/2017) the bees collected from the entrance of 22 free-living colonies were screened for Nosema spores using light microscopy and a visual check of bee bodies for varroa. The numbers of bees tested per colony ranged from two to 46. In 2017/2018 the sugar shake method (Lee et al 2010) was applied to bees sampled from nine colonies (numbers of bees ranged from 30-80).

DNA extraction, PCR and sequencing

DNA was extracted from the two hind legs of worker bees using the E.Z.N.A. Forensics DNA extraction kit (Omega Bio-Tek). Mitochondrial DNA consisting of the 3' end of the $tRNA^{leu}$ gene, and the 5' end of the COII subunit gene was amplified using E2 (5'-GGCAGAATAAGTGCATTG-3') and H2 (5'-CAATATCATTGATGACC-3') primers (Garnery *et al.* 1998) with Illustra PuReTaq Ready-To-Go PCR Beads (GE Healthcare). PCR reactions included an initial denaturation of 5 mins at 95 °C, followed by 35 cycles of 94 °C for 45 secs, 45 °C for 43 secs and 62 °C for 2 mins with a final extension of 20 min at 65 °C (Garnery *et al.* 1993). PCR products were purified using a GeneJet PCR Purification kit (Thermo scientific). Sequencing of PCR fragments was by Sanger sequencing at LGC Genomics, Germany. The sequences were manually assessed against their chromatographs in MEGA7 (Kumar *et al.* 2016) before each one was imported into a multiple alignment. All unique sequences have been deposited into GenBank (See Hassett *et al*. 2018).

Mitochondrial DNA sequence analysis

Ninety nine sequences of satisfactory quality were generated from 49 of the 76 free-living colonies (Table 1). Due to spurious base calling at the 5' and 3' ends, sequences were trimmed to start at base 621, the TTAATAAA motif, 5'end of the P element, and to end at base 1547, the end of a TTTTTTTT motif, (*sensu* Cornuet *et al.* 1991). Base calling at sites 980 and 1129 (A/T and A/G, respectively) was determined to be unreliable, due to differences between sister bees and even repeat sequences from the same bee, therefore both sites were removed as were identical sequences between sister bees. Initial alignments contained all available European *A. m. mellifera* sequences (https://www.ncbi.nlm.nih.gov/genbank/) along with representatives of other *A. mellifera* sub-species, and 156 sequences from the managed cohort in Ireland (Hassett *et al* 2018). Relationships were examined using networks as

implemented in TCS 1.21 (Clement *et al.* 2000). Subsequently, given the large number of sequences, the presence of very distinct mitotypes and unconnected networks resulting from initial analyses, a smaller alignment was created of more closely related sequences containing the sequences from Irish free-living bees along with 14 *A. m. mellifera* sequences from protected black bee populations in Europe (Pinto *et al.* 2014). This final alignment (Supplementary data file *feral52.fas)* was analysed in TCS 1.21 (Clement *et al.* 2000) and a network produced. Given the large influence of the duplication events, that have led to the presence of extra Q elements in some bees, on the resulting relationships between the sequences, an additional alignment was analysed that included just the first position of each Q element as well as any point mutations evident within a Q element.

Microsatellite analysis

DNA from 59 free-living colonies (Table 1) was sent to Ecogenics, Switzerland, for genotyping using a twelve microsatellite panel (A273, A43, Ac306, Ap33, B24, Ap226, A76-2p, A007, Ap001, A28, Ap289 and A29), which had previously been chosen for their informative value in the analysis of *A. m. mellifera* (Hassett *et al.* 2018). Equivalent data were included from reference populations of *A. m. ligustica* from Italy (n = 55) and *A. m. carnica* from Austria (n = 182) and Slovenia (n = 21), *A. m. mellifera* from Sweden (n = 10), France (n = 24), Norway (n = 18) and Switzerland (n = 22). The dataset analysed was completed by the addition of data from 171 managed bees from Ireland (Hassett *et al.* 2018). Bayesian analysis and visualisation of population assignment between C and M lineages was conducted in STRUCTURE V2.3.4 (Pritchard *et al.* 2000) using the admixture and correlated allele frequency models with the unsupervised option. A total of 750,000 Markov chain Monte Carlo (MCMC) iterations after an initial burn-in of 250,000 were performed for 20 iterations of each of K = 1 to 6. The optimal value of K (Evanno *et al.* 2005) was calculated using the CLUMPAK (Kopelman *et al.* 2015) online calculator. Nine populations were designated prior to analysis; seven based on the reference populations as above, and two for the Irish population divided between managed and free-living colonies (1: Italy ligustica, 2: Austrian carnica, 3: Slovenian carnica 4: Swedish mellifera, 5: French mellifera, 6: Norwegian mellifera, 7: Swiss mellifera 8: Irish managed and 9: Irish free-living). The threshold used for full assignment to a particular population was a Q-value ≥ 0.900 (Vaha and Primmer 2006). The populations were arranged into four groups: European C-lineage (populations. 1, 2 and 3), European mellifera (populations. 4, 5, 6 and 7),

Irish managed (population 8) and Irish free-living (population 9). Using these groups an analysis of molecular variance (AMOVA) was conducted using Arlequin V3.5.2.2 (Excoffier and Lischer 2010).

Table 1. Details of all free-living honey bee colonies sampled including information on their indicated lineage (M versus C) via mitochondrial, microsatellite and SNPS data. Msats= microsatellite. Colonies highlighted in grey are those that have results from all three datatypes. The three colonies in bold are those where one or more data type indicate introgression of C into M lineage.

	mtDNA		Msats		SNPs	
Colony ID	**♯bees**	**Lineage**	**♯ bees**	**Prop M**	**♯ Bees**	**Prop M**
F1T	nd	nd	nd	nd	1	0.9779
F2L	1	M	2	0.997-8	1	0.9888
F3CE	nd	nd	nd	nd	9	0.999
F4L	2	M	3	0.996-8	1	1
F5LA	nd	nd	3	0.9980	1	0.9905
F5LB	1	M	nd	nd	nd	nd
F6L	1	M	2	0.9960	nd	nd
F7L	1	M	1	0.9970	nd	nd
F8KEA	nd	nd	2	0.992-5	1	0.9905
F8KEB	1	M	1	0.9980	1	0.9799
Fl9DA	1	M	1	0.9980	1	1
Fl9DB	1	M	1	0.9810	1	0.976
F10G	1	M	1	0.9970	1	1
F11R	1	M	nd	nd	nd	nd
F12R	1	M	2	0.9980	1	0.9798
F13R	1	M	2	0.989-0.993	1	1
F14G	2	M	1	0.9980	9	1
F15G	2	M	1	0.9890	nd	nd
F16G	1	M	1	0.9970	9	0.9833
F17G	nd	nd	nd	nd	nd	nd
F18G	1	M	nd	nd	1	0.9683
F19D	1	M	nd	nd	1	1
F20CE	1	M	2	0.9980	9	0.9839
F21C	**5**	**M**	**4**	**0.607-0.998**	**1**	**0.7281**
F22	1	M	1	0.9960	nd	nd
F23	2	nd	nd	nd	nd	nd
F24G	2	M	nd	nd	nd	nd
F25G	1	M	nd	nd	nd	nd
F26g	nd	nd	nd	nd	nd	nd
F27G	2	M	2	0.9980	9	1
F28G	1		2	0.9980	1	1
F29G	7	M	5	0.933-997	8	1
F30G	1	M	1	0.9980	9	1
F31G	2	M	1	0.9900	nd	nd
F32WX	nd	nd	nd	nd	1	0.9744
F33KY	nd	nd	1	0.9980	1	1

F34W	2	M	1	0.9960	nd	nd
F35CW	1	M	nd	nd	nd	nd
F36C	2	M	1	0.9970	nd	nd
F37C	0		1	0.9910	1	0.9903
F38C	2	M	1	0.9970	nd	nd
F39G	nd	nd	nd	nd	nd	nd
F40G	nd	nd	nd	nd	nd	nd
F41G	3	M	2	0.993-4	9	0.9999
F42G	nd	nd	1	0.9970	nd	nd
F43G	1	M	1	0.9980	nd	nd
F44G	7	M	5	**0.893-0.997**	nd	nd
F45G	2		2	0.9980	9	0.9884
F46G	nd	nd	nd	nd	nd	nd
F47G	3	M	2	0.989-0.9980	9	1
F48G	8	M	4	0.997-0.998	9	1
F49R	1	M	1	0.9940	nd	nd
F50R	2	M	1	0.9890	nd	nd
F51MO	1	M	1	0.9960	1	0.994
F52D	1	M	1	0.9860	1	0.9672
F53Lo			1	0.9970	1	0.9853
F54MN	2	M	2	0.986-0.997	1	0.9429
F55MN	2	M	2	**0.821, 0.988**	1	0.9665
F56LS	2	M	1	0.9940	1	0.9684
F57OY	2	M	1	0.997-8	1	0.9908
F58C	nd	nd	1	0.9960	nd	nd
F59L	nd	nd	1	0.9980	nd	nd
F60L	nd	nd	1	0.9920	nd	nd
F61L	5	M	4	0.995-8	nd	nd
F62L	nd	nd	1	0.9940	nd	nd
F63WX	nd	nd	1	0.9980	nd	nd
F64CE	nd	nd	1	0.9740	nd	nd
F65OY	nd	nd	1	0.9970	nd	nd
F66G	nd	nd	1	0.9160	nd	nd
F67G	1	M	1	0.9730	nd	nd
F68G	1	M	1	0.9980	nd	nd
F69WW	nd	nd	1	0.9970	nd	nd
F70CE	nd	nd	1	0.9850	nd	nd
F71G	1	M	nd	nd	nd	nd
F72G	1	M	2	0.9970	nd	nd
F73G	1	M	1	0.9970	nd	nd
	99 bees		**94 bees**		**123 bees**	
76 Colonies	**53 colonies**		**59 colonies**		**36 colonies**	

Note: # = number of, Msat=microsatellite, SNP=Single Nucleotide Polymorphism, Prob M=probability of individual bee being part of M lineage. Nd=Not done. Colonies highlighted in grey are those that have results from all three types of data. The three colonies in bold are those where one or more data type indicate introgression of C into M lineage.

Single Nucleotide Polymorphisms (SNPs) analysis

DNA from 127 free-living bees representing 39 colonies were diluted to 10–15 ng/µl and sent to Instituto Gulbenkian de Ciência (Portugal) for genotyping using the Agena BioScience iPLEX chemistry and the MassARRAY® MALDI-TOF platform (Gabriel *et al* 2009) using a highly informative 127 SNP assay, designed for reliable introgression estimation of C-into M-lineage (Henriques *et al*. 2018). After quality control to identify SNPs with missing data (>20%), data from 36 colonies were kept in the final analysis (Table 1). For most colonies one bee was tested, for 11 colonies nine bees were included and for one colony eight bees were included (Table 1). Membership proportions (Q-values) were estimated using ADMIXTURE V1.23 (Alexander *et al* 2009) for K = 2 with 20 independent runs of 10,000 iterations. The convergence between iterations was examined by comparing log-likelihood scores (LLS) using the default termination criteria set to stop when LLS increases by <0.0001 between iterations. A total of 36 M-lineage and 36 C-lineage individuals were used as a reference population. CLUMPAK was used to summarise and visualise the Q-values. An arbitrary threshold of Q-value ≥0.950 was considered full assignment to either lineage.

RESULTS

Location, health, and survival of free-living colonies

Between November 2015 and November 2018, a total of 209 reports of putatively free-living bee colonies were received of which 7.2% (15) were identified as bumble bees, solitary bees, or wasps. A further 5.7% (12) consisted of captured swarms with unconfirmed provenance. Colonies reported in Spring 2016, before swarming had begun, were assumed to have been present since late Autumn 2015. The reported habitats of all remaining colonies (n=182) consisted primarily of cavities in buildings (68%) where mainly the roof space was occupied by the colony whereas trees formed the second most utilised habitat at 10% (Table 2). Of the 76 colonies that were monitored for survival from Autumn 2015 to Spring 2019, the survival reports on 16 colonies were considered ambiguous and removed from further study, 21 colonies (27.63%) survived for 2-2.5 years and 22 (28.95%) survived three or more years (Figure 2). Colony deaths were primarily non-survival over winter or in early spring; the exact causes were unknown. External causes of death were extermination by the home owner and one case of predation by Pine marten (*Martes martes*). Of the 22 colonies screened for varroa

and nosema in 2016/2017 one varroa mite was found from one colony and no *Nosema* spp. spores were detected. In 2017/2018 again, a single varroa mite was detected in one colony.

Table 2: Residency periods (as reported by citizen scientists) and range for 182 honey bee colonies by habitat type.

Habitat	No.(%)of colonies	Mean residency (Yrs)	Residency range (Yrs)
Buildings	124 (68%)	4.6	1 to 40
Free standing walls	7 (4%)	6.0	1 to 30
Trees	18 (10%)	5.9	1 to 40
Other	33 (18%)	1.2	1 to 5
All types	**182**	**4.2**	**1 to 40**

"Other" includes, *inter alia*, graves, a statue, a cattle grid and a bird nest box.

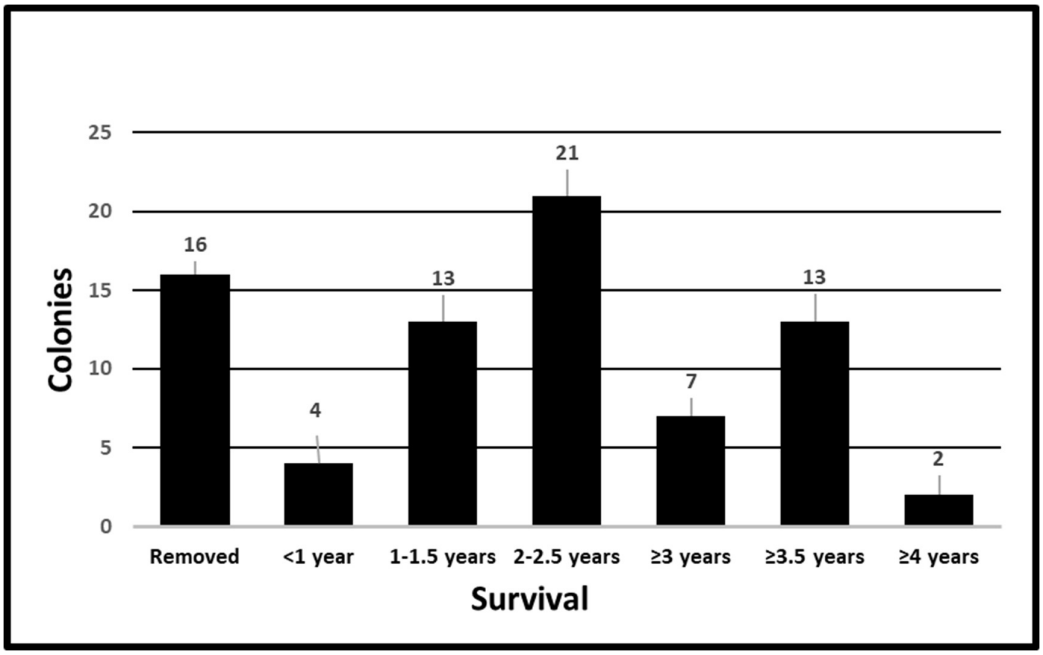


Figure 2: Survival of colonies (n=76) monitored between Autumn 2015 and Spring 2019. Sixteen colonies were removed from survival data due to incomplete records or ambiguity regarding survival, as reported by the custodian (citizen scientist). Twenty-two colonies survived at least three winters of which two survived four. In some cases, custodians reported that locations have been housing colonies for decades.

Population structure from mitochondrial data

A total of 99 COI-COII mitochondrial sequences were generated from free-living colonies and 100% were *A. m. mellifera* mitotypes, containing the P element (Cornuet *et al.* 1991). After pruning to remove identical sequences from the same colony, 52 sequences remained: 37 PQQ, 11 PQQQ and four PQQQQ mitotypes (Figure 3). When the entire Q elements, alongside the intergenic P sequences, were included in the TCS analyses (Clement *et al.* 2000), separate networks were formed for sequences that contained different numbers of duplicated Q elements (PQQ, PQQQ or PQQQQ). When the Q elements were represented only by one base pair each, plus the seven sites where variation occurred within them, three clusters with PQQ, PQQQ or PQQQQ were still evident in the network, but they were now connected (Figure 3) except for a single mitotype in the PQQ cluster (F36C056) and a group of three in the PQQQ cluster (F28G267, F35CW055 and F51MO096) which were distinct.

Trimming of the sequences lead to no differentiation between the M4d and M4e, and the M4a and M4m mitotypes. As shown in Figure 3, Just over half of colonies (n=27, 52%) yielded sequences that were identical to M4e and M4d reference mitotypes and two of the Irish free-living bee sequences were identical to M4f while the rest of the mitochondrial sequences (23, 44%) were distinct from any available European sequence. In comparison between the data from free-living bees and managed bees sampled to date from Ireland, seven variants were found only in the free-living bees. Four of these sequences (F31G50, F45G085, F13R036, and F43G080/FKe017) had variations represented by single indels or point mutations. Three variants (F6L011, F36C056 and F28G267/F35CW055/F51MO096) had multi-base deletions of between 6 bp and 10 bp and form a distinct PQQQ cluster (Figure 3). A fourth variant from two colonies (F18G104/F47G256) with a significant deletion (6 bp from sites 17 to 22) was also found in the Irish managed population but not in any available data from elsewhere in Europe as were the remaining variants shown on Figure 3.

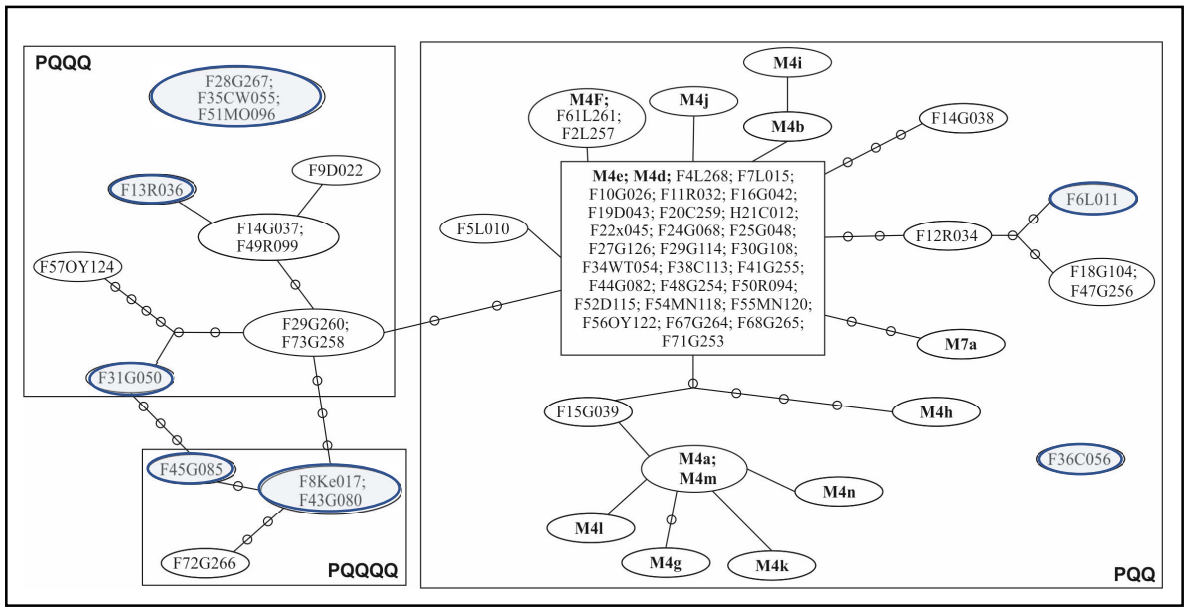


Figure 3: Statistical parsimony network (TCS v1.21) of the 19 different mitotypes from 99 mtDNA sequences generated from Irish free-living bees. These have been included with European mitotypes downloaded from GenBank. Each circle (o) on a branch represents a single indel or point mutation. Branch lengths are not representative of distance. The shaded ovals contain the variants that were only found in the free-living population, none were seen in the managed population sampled so far.

Population differentiation from microsatellite and SNP data

Both microsatellite and SNP data (Figures 4 & 5 respectively) indicate clear structure between the C-lineage bees and the M-lineage bees for K=2. Irish free-living bees showed a high degree of purity in both analyses. In the microsatellite analysis, Irish managed bees (Figure 4, population 8) 167 (97.7%) of the 171 bees were also indicated as pure M-lineage. Those that fell below the assignment threshold included individuals from two colonies considered to be either Buckfast or hybrids,. For the free-living bees using these data three colonies (5%) fell below Q-value ≥ 0.900 indicating that they cannot be assigned clearly to a particular lineage. F21Ce was located in County Clare in an area known to contain beekeepers that keep Buckfast. This colony contained a mixture of bees with different levels of putative purity with some bees being assigned to M lineage with Q value of 0.998 while one bee had a Q value of 0.61. F44G was collected in the walls of a castle in east Galway and again showed bees with mixed lineage, some assigned to M lineage with Q value of 0.997 and others dropping below the 0.9 cut off for confident lineage assignment. Similarly, F55MN, collected in the roof of an old cottage in County Monaghan, showed one bee that could be assigned to M lineage with confidence while the other bee could not (Table 1). SNPs data identified only one colony falling below a Q value threshold of 0.9, and this was F21Ce, also identified by microsatellites

as showing introgression. While lineage assignment was not tested for F44G via the SNPs approach, the bee from colony F55Mn had a Q value of 0.97 while a different colony also from Monaghan showed a Q value between 0.9 and 0.95 (F54MN, Table 1). Where SNPs data were returned for 8/9 bees per colony, all bees from all 12 colonies could be assigned to the M lineage clearly. Unfortunately, F21Ce, F44G and F55Mn were not included in that experiment.

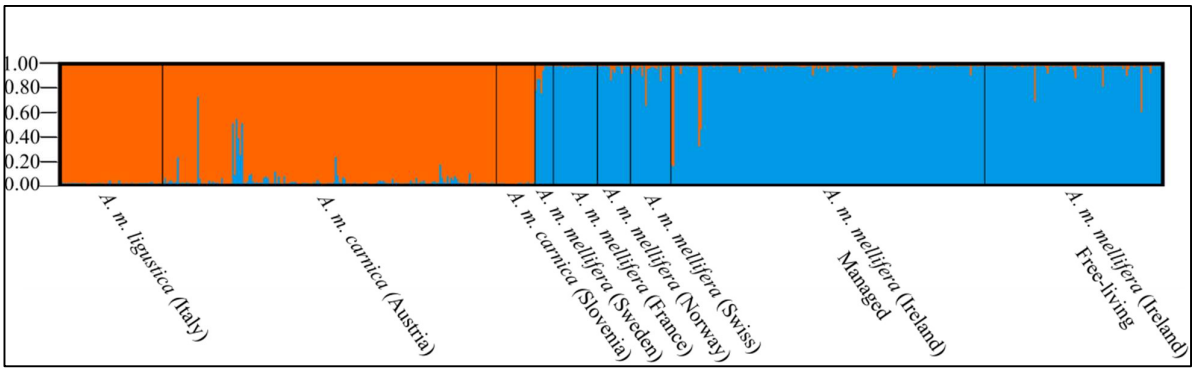


Figure 4: Structure K=2 assignments of 598 individuals from 9 putative populations using twelve microsatellites (1: Italy ligustica, 2: Austrian carnica, 3: Slovenian carnica 4: Swedish mellifera, 5: French mellifera, 6: Norwegian mellifera, 7: Swiss mellifera 8: Irish managed and 9: Irish free-living) including 95 individual honey bees from 50 free-living colonies. Each vertical bar represents an individual bee with assignment apportioned between orange (C-lineage) and blue (M-lineage). Assignment values are from 0.000 to 1.000. Black vertical lines separate the populations.

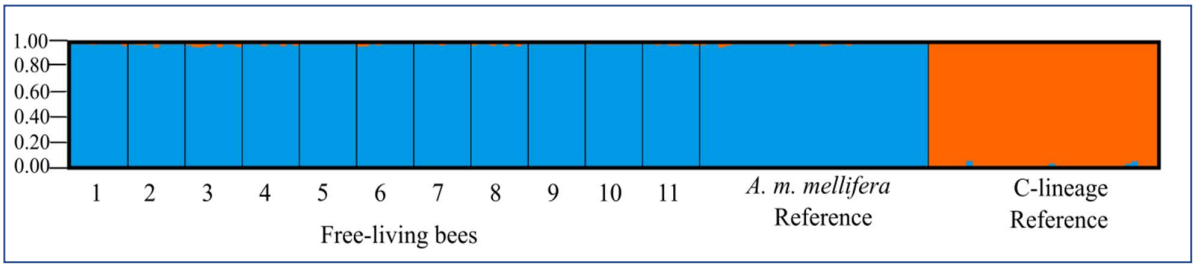


Figure 5: K=2 ADMIXTURE assignments of SNPs genotyping of 99 free-living bees over 11 different colonies (9 per colony) along with 36 M-lineage *A. m. mellifera* and 36 C-lineage reference bees. Q value assignments are from 0.000 to 1.000 with a value ≥0.950 denoting full assignment to a lineage. Each vertical bar represents an individual bee with assignment apportioned between Blue = M-lineage and Orange = C-lineage. Black vertical lines separate the colonies/groups.

AMOVA indicated that while a proportion (38.59%) of the genetic variation existed among the M and C groups, greater variation (59.62%) occurred within the populations. No real distinction was evident between the managed and free-living honey bee cohorts in Ireland. The low pairwise F_{ST} value, 0.013, between the Irish managed and free-living cohorts, was similar but lower than that between the Austrian and Slovenian *A. m. carnica* populations (0.017) which border each other (Figure 6).

Pop.	***Italian***	***Austrian***	***Slovenian***	**Swedish**	**French**	**Norwegian**	**Swiss**	**Irish man.**	**Irish free**
Italian	0.00000								
Austrian	0.33430	0.00000							
Slovenian	0.41583	0.01717	0.00000						
Swedish	0.47367	0.34234	0.45152	0.00000					
French	0.57277	0.41541	0.55052	0.17103	0.00000				
Norwegian	0.53278	0.39605	0.51355	0.09895	0.05124	0.00000			
Swiss	0.55013	0.38989	0.51490	0.14800	0.02248	0.03596	0.00000		
Irish man.	0.53454	0.41574	0.51098	0.16769	0.03143	0.05843	0.05660	0.00000	
Irish free	0.54576	0.40608	0.51351	0.16508	0.05418	0.06556	0.05966	0.01334	0.00000

Figure 6: Population pairwise Fst matrix by distance method of microsatellite alleles between the 9 honey bee populations: Italian ligustica, Austrian carnica, Slovenian carnica, Swedish mellifera, French mellifera, Norwegian mellifera, Swiss mellifera, Irish managed and Irish free-living.

DISCUSSION

This study provides the first fully documented evidence, since the discovery of *V. destructor* in Ireland in 1998, that free-living, untreated, honey bees exist in the country and that they primarily use the cavities located in new and historic buildings to house their colonies. Anecdotal evidence had already indicated the likelihood of the existence of free-living colonies that were capable of longer-term survival without human intervention and the data collected so far strengthen this. While the majority of colonies have been directly monitored as part of this study for a period of only three years, some of these colonies had already been in place for a period prior to the study commencing. Continued monitoring combined with a molecular assay to assess the persistence of the original queen lineage more accurately is needed to fully elucidate the longevity of these lineages.

Contrary to our findings where 68% of colonies were located in buildings, in the UK there was no significant difference between the use of trees, houses and walls for colony sites (Thompson, 2012). In Ireland approximately 11% land cover is woodland, the majority of which is commercially grown coniferous species (D.A.F.M. 2018), which are usually thinned at 15 years of growth and felled by 30 years. A high turnover of trees in managed forests combined with loss of mature deciduous woodland is likely to produce a low density of trees with cavities of sufficient size for colonisation by honey bees. Conversely, although the UK has a similar relative woodland cover (13%) it consists of roughly even areas of coniferous and broadleaved woodland, which creates a higher age profile than that found in Ireland (Forest Research UK, 2018) and this may help explain the relatively greater use of trees by UK colonies. Colonies in house cavities obtain the benefits of a long lasting, insulated space, giving them the time needed to expand both individual colonies and the dynasty of the queen, an arrangement which may be a benefit over tree cavities.

Evidence now hints at colony survival longevity, past the stage where Varroosis related death would be expected (Korpela *et al.* 1992). In interpreting the possible mechanisms behind the survival of free-living colonies, their cryptic nature not only makes it difficult to locate colonies but also to directly test for levels of varroa infestation using techniques such as daily mite drops, alcohol rolls or sugar shaker. The controlled observation of collected swarms known to originate from survivor colonies may help elucidate these mechanisms. Additionally,

providing bee boxes for swarms to occupy, that were accessible for sampling of the colony would facilitate pathogen screening and may increase our knowledge of how free-living colonies fit into the epidemiology of varroa and other honey bee diseases. There is a concern amongst beekeepers that disease prevalence and loads undergo expansion in the unmanaged honey bee population and that this places managed colonies at greater risk. Greater access to a sample portion of the free-living population could help determine the validity or otherwise of this concern. Despite the current difficulty in sampling free-living colonies for the presence of parasites and pathogens, where possible, sampled bees were screened for varroa and Nosema. The bees sampled were only from the entrance, which excluded testing the nurse bees which are more attractive to varroa (Del piccolo *et al.* 2010) and would normally form part of a test such as the sugar shaker assay. However, colonies were sampled in early Autumn when varroa becomes more phoretic and might be more common in foragers. Although it may seem highly unlikely that free-living colonies would be free of such parasites the limited results to date provide some optimism that loads may not be very high.

Over 50% of the mitotypes identified in the Irish free-living colonies were identical to Dutch mitotypes. This is clearly associated with the period at the beginning of the 20th century after the outbreak of The Isle of Wight disease in Ireland when there was a large importation of replacement honey bees from Holland. Remarkably none of the free-living bees showed mitotypes from elsewhere in Europe (e.g. not identical to any of the sequences from Rortais *et al* 2010). Hassett *et al* (2018) revealed two Irish bees with identical mitochondrial sequences to a French mitotype and one bee showed identity to bees from Colonsay. These bees were from the managed cohort and together the results so far indicate little influence of European *A. m. mellifera* in Ireland apart from a major influence from the importations from Holland. Most of the mitochondrial variants found in the free-living bees could also be found within the managed Irish population (Hassett *et al.* 2018), meaning Ireland has a free-living population that appears fundamentally undifferentiated from the managed one, also confirmed by AMOVA of microsatellite data. That said, of the 29 new variants described in Hassett *et al* (2018), seven were found exclusively in the free-living bees, which might indicate unique genetic variation present in free-living bees. However, this is a relatively small study and additional sampling may identify the putatively exclusive free-living variants in the managed population.

Interestingly, the retention of clusters that are essentially defined by the number of Q elements present, once the majority of the Q element sequence has been excluded from inclusion in the analyses, may provide some evidence that each duplication occurs as a single event such as the AT-rich homologous motifs suggested by (Cornuet *et al.* 1991). If the duplication event is a synapomorphy then the Q element architecture represents separate M lineages, creating a problem with the current naming system for mitotypes. Putative mitochondrial lineages as evident in Figure 3 and in Hassett *et al* (2018) will need to be confirmed with other data.

From a conservation viewpoint, it is important that both microsatellite and SNP data indicate that, in keeping with the Irish population as a whole, the free-living population sampled consists mostly of bees that can be assigned to *A. m. mellifera* with high confidence. Using SNPs, comparable levels of purity were returned whether one or nine worker bees was used to represent the colony reinforcing the argument of (Henriques *et al.* 2018) that genotyping one single bee per colony is sufficient to indicate purity. However, with microsatellite data generated from a widespread sample of the Irish population different lineages were identified in a single colony when 2-5 bees were included for analysis. It is clear that families of bees with different fathers do exist in some colonies particularly in places where the queen may not have enough drones of the same subspecies with which to mate. We would recommend the processing of more than 2 bees in locations where the presence of non-native subspecies and hybrids are suspected.

In periods as short as ten years, experimental survivorship tests in isolated areas have indicated that a balanced host-parasite relationship can develop in colonies from a small population (Fries *et al.* 2006, Le Conte *et al.* 2007). The high rate of genetic recombination in honey bees (Beye *et al.* 2006) may produce a sufficiently diverse population from only a few survivors to allow genes linked to resistance/tolerance mechanisms to quickly proliferate in a small population. However, on a broader scale some degree of mating isolation may be essential since the panmictic mating of honey bees may hinder these genes from becoming fixed in the population (van Alphen, 2018). Under natural conditions, colony density can return to the levels present before the arrival of varroa (Mikheyev *et al.* 2015, Seeley 2007) and given the low levels of commercialised beekeeping in Ireland, combined with the existence of a putatively large free-living population, it seems plausible that colonies with

varroa resistance/tolerance mechanisms have emerged in Ireland over the 20 years since varroa was first discovered. Colonies with traits that allow survival in the presence of varroa probably existed in Irish apiaries (as well as in the wild population), however chemical treatment against varroa would not have allowed these resistant/tolerant colonies to stand out from those that were susceptible to varroa. These genetic goldmines could have swarmed into the surrounding areas providing the resources on which natural selection could act and augmenting any free-living bees managing to survive with similar or different varroa resistant/tolerant traits. If this can be substantiated, the lack of differentiation between the managed and free-living populations may not be considered a negative condition for the discovery of varroa resistant/tolerant genotypes.

Given the interactions between genotype and environment that have been clearly shown in honey bees (Buechler *et al.* 2014, Meixner *et al.* 2014) and the fact that local ecotypes seem to do better than bees translocated from different microclimates (Costa *et al.* 2012), there is also a need to protect local adaptations. Research on free-living bees in Ireland now requires expansion to allow an accurate indication of colony density on a national scale, along with observation of survivorship and associated mechanisms, as well as characterisation of local strains. We believe that the results presented here, combined with those of Hassett *et al*. (2018) and the observations of beekeepers, require immediate application of the precautionary principle of conservation practice (Finnoff *et al.* 2007) for the protection of Ireland's locally adapted free-living honey bee population. We join with other researchers in requesting legal protection for local adaptations in *A. mellifera* (Fontana *et al.* 2018) through stricter control on the movement of live bees and the banning of imports. Inter-country movement of *A. m. mellifera* even where it is the indigenous sub-species needs to be given careful consideration in each case to avoid out-breeding locally adapted gene complexes. However, having been the welcome recipient of Dutch honey bees following the collapse of Irish beekeeping at the beginning of the 20th century, Ireland may yet be able to return the favour by returning bees of Dutch mitotypes home to the Netherlands from a free-living population that may have developed some degree of varroa resistance, while at the same time increasing the commercial viability of its beekeeping practice.

References

Alexander, D. H., Novembre, J. and Lange, K. (2009) 'Fast model-based estimation of ancestry in unrelated individuals', *Genome Research,* 19(9), 1655-1664.

Alice Pinto, M., Henriques, D., Chavez-Galarza, J., Kryger, P., Garnery, L., van der Zee, R., Dahle, B., Soland-Reckeweg, G., de la Rua, P., Dall'Olio, R., Carreck, N. L. and Johnston, J. S. (2014) 'Genetic integrity of the Dark European honey bee (*Apis mellifera mellifera*) from protected populations: a genome-wide assessment using SNPs and mtDNA sequence data', *Journal of Apicultural Research,* 53(2), 269-278.

Anderson, D. L. and Trueman, J. W. H. (2000) '*Varroa jacobsoni* (Acari : Varroidae) is more than one species', *Experimental and Applied Acarology,* 24(3), 165-189.

Beye, M., Gattermeier, I., Hasselmann, M., Gempe, T., Schioett, M., Baines, J. F., Schlipalius, D., Mougel, F., Emore, C., Rueppell, O., Sirvio, A., Guzman-Novoa, E., Hunt, G., Solignac, M. and Page, R. E. (2006) 'Exceptionally high levels of recombination across the honey bee genome', *Genome Research,* 16(11), 1339-1344.

Buechler, R., Costa, C., Hatjina, F., Andonov, S., Meixner, M. D., Le Conte, Y., Uzunov, A., Berg, S., Bienkowska, M., Bouga, M., Drazic, M., Dyrba, W., Kryger, P., Panasiuk, B., Pechhacker, H., Petrov, P., Kezic, N., Korpela, S. and Wilde, J. (2014) 'The influence of genetic origin and its interaction with environmental effects on the survival of *Apis mellifera* L. colonies in Europe', *Journal of Apicultural Research,* 53(2), 205-214.

Chauzat, M. P., Cauquil, L., Roy, L., Franco, S., Hendrikx, P. and Ribiere-Chabert, M. (2013) 'Demographics of the European Apicultural Industry', *Plos One,* 8(11).

Clement, M., Posada, D. and Crandall, K. A. (2000) 'TCS: a computer program to estimate gene genealogies', *Molecular Ecology,* 9(10), 1657-1659.

Cornuet, J. M., Garnery, L. and Solignac, M. (1991) 'Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA', *Genetics,* 128(2), 393-403.

Costa, C., Lodesani, M. and Bienefeld, K. (2012) 'Differences in colony phenotypes across different origins and locations: evidence for genotype by environment interactions in the Italian honeybee (*Apis mellifera ligustica*)?', *Apidologie,* 43(6), 634-642.

De la Rua, P., Jaffe, R., Dall'Olio, R., Munoz, I. and Serrano, J. (2009) 'Biodiversity, conservation and current threats to European honeybees', *Apidologie,* 40(3), 263-284.

D.A.F.M. (Department of Agriculture Fisheries and Marine) website www.agriculture.gov.ie/forestservice/ (Accessed 23rd Nov. 2018)

Ellis, J. S., Soland-Reckeweg, G., Buswell, V. G., Huml, J. V., Brown, A. and Knight, M. E. (2018) 'Introgression in native populations of *Apis mellifera mellifera* L: implications for conservation', *Journal of Insect Conservation,* 22(3-4), 377-390.

Evanno, G., Regnaut, S. and Goudet, J. (2005) 'Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study', *Molecular Ecology,* 14(8), 2611-2620.

Excoffier, L. and Lischer, H. E. L. (2010) 'Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows', *Molecular Ecology Resources,* 10(3), 564-567.

Finnoff, D., Shogren, J. F., Leung, B. and Lodge, D. (2007) 'Take a risk: Preferring prevention over control of biological invaders', *Ecological Economics,* 62(2), 216-222.

Fontana *et al* (2018) Appeal for biodiversity protection of native honey bee subspecies of *Apis mellifera* in Italy (San Michele all'Adige declaration) *Bulletin of Insectology* 71 (2) 257-271

Forest Research UK (2018) www.forestresearch.gov.uk/tools-and-resources/national-forest-inventory/ Accessed 26 Nov. 2018

Fries, I., Imdorf, A. and Rosenkranz, P. (2006) 'Survival of mite infested (varroa destructor) honey bee (*Apis mellifera*) colonies in a Nordic climate', *Apidologie,* 37(5), 564-570.

Gabriel, S., *et al* (2009) 'SNP genotyping using the Sequenom MassARRAY iPLEX platform', *Current Protocols in Human Genetics,* 60:2.12. 11-12.12. 18

Hassett, J., Browne, K. A., McCormack, G. P., Moore, E., Soland, G., Geary, M. and Native Irish Honey Bee, S. (2018) 'A significant pure population of the dark European honey bee (*Apis mellifera mellifera*) remains in Ireland', *Journal of Apicultural Research,* 57(3), 337-350.

Henriques, D., Browne, K. A., Barnett, M. W., Parejo, M., Kryger, P., Freeman, T. C., Munoz, I., Garnery, L., Highet, F., Jonhston, J. S., McCormack, G. P. and Pinto, M. A. (2018) 'High sample throughput genotyping for estimating C-lineage introgression in the dark honeybee: an accurate and cost-effective SNP-based tool', *Scientific Reports,* 8.

Jaffe, R., Dietemann, V., Allsopp, M. H., Costa, C., Crewe, R. M., Dall'olio, R., de la Rua, P., El-Niweiri, M. A. A., Fries, I., Kezic, N., Meusel, M. S., Paxton, R. J., Shaibi, T., Stolle, E. and Moritz, R. F. A. (2010) 'Estimating the Density of Honeybee Colonies across Their Natural Range to Fill the Gap in Pollinator Decline Censuses', *Conservation Biology,* 24(2), 583-593.

Jensen, A. B., Palmer, K. A., Boomsma, J. J. and Pedersen, B. V. (2005) 'Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honeybee, *Apis mellifera mellifera*, in northwest Europe', *Molecular Ecology,* 14(1), 93-106.

Kohl, P. and Rutschmann, B. (2018) 'The neglected bee trees: European beech forests as a home for feral honey bee colonies', PeerJ, 6:e4602

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. and Mayrose, I. (2015) 'Clumpak: a program for identifying clustering modes and packaging population structure inferences across K', *Molecular Ecology Resources,* 15(5), 1179-1191.

Korpela, S., Aarhus, A., Fries, I. and Hansen, H. (1992) 'varroa jacobsoni OUD in cold climates: Population growth, winter mortality and influence on survival of honey bee colonies', *Journal of Apicultural Research,* 31(3-4), 157-164.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets', *Molecular Biology and Evolution,* 33(7), 1870-1874.

Le Conte, Y., De Vaublanc, G., Crauser, D., Jeanne, F., Rousselle, J.-C. and Becard, J.-M. (2007) 'Honey bee colonies that have survived varroa destructor', *Apidologie,* 38(6), 566-572.

Lee, K. V., Moon, R. D., Burkness, E. C., Hutchison, W. D. and Spivak, M. (2010) 'Practical Sampling Plans for *Varroa destructor* (Acari: Varroidae) in *Apis mellifera* (Hymenoptera: Apidae) Colonies and Apiaries', *Journal of Economic Entomology,* 103(4), 1039-1050.

McMullan, J. (2018) 'Adaptation in Honey Bee (*Apis mellifera*) Colonies Exhibiting Tolerance to *Varroa destructor* in Ireland', *Bee World,* 95(2), 39-43.

Meixner, M. D., Buechler, R., Costa, C., Francis, R. M., Hatjina, F., Kryger, P., Uzunov, A. and Carreck, N. L. (2014) 'Honey bee genotypes and the environment', *Journal of Apicultural Research,* 53(2), 183-187.

Mikheyev, A. S., Tin, M. M. Y., Arora, J. and Seeley, T. D. (2015) 'Museum samples reveal rapid evolution by wild honey bees exposed to a novel parasite', *Nature Communications,* 6, 7991

Moritz, R. F. A., Bernhard Kraus, F., Kryger, P. and Crewe, R. M. (2007) 'The size of wild honeybee populations (*Apis mellifera*) and its implications for the conservation of honeybees', *Journal of Insect Conservation,* 11(4), 391-397.

Moritz, R. F. A., Hartel, S. and Neumann, P. (2005) 'Global invasions of the western honeybee (*Apis mellifera*) and the consequences for biodiversity', *Ecoscience,* 12(3), 289-301.

Parejo, M., Henriques, D., Pinto, M. A., Soland-Reckeweg, G. and Neuditschko, M. (2018) 'Empirical comparison of microsatellite and SNP markers to estimate introgression in *Apis mellifera mellifera*', *Journal of Apicultural Research,* 57(4), 504-506.

Pritchard, D. (2018) 'Geordie bees bite the mite', SICAMM conference, Finland, 2018

Pritchard, J. K., Stephens, M. and Donnelly, P. (2000) 'Inference of population structure using multilocus genotype data', *Genetics,* 155(2), 945-959.

Randi, E. (2008) 'Detecting hybridization between wild species and their domesticated relatives', *Molecular Ecology,* 17(1), 285-293.

Rortais, A., Arnold, G., Alburaki, M., Legout, H. and Garnery, L. (2011) 'Review of the DraI COI-COII test for the conservation of the black honeybee (*Apis mellifera mellifera*)', *Conservation Genetics Resources,* 3(2), 383-391.

Seeley, T. D. (2007) 'Honey bees of the Arnot Forest: a population of feral colonies persisting with *Varroa destructor* in the northeastern United States', *Apidologie,* 38(1), 19-29.

Soland-Reckeweg, G. (2006) Genetic differentiation and hybridization in the honeybee (*Apis mellifera* L .) in Switzerland . Thesis, University of Bern.

Swindon honeybee conservation group (2019) website (Accessed 5th August 2019) http://www.swindonhoneybeeconservation.org.uk/research/

Thompson, C.E. (2012) The health and status of the feral honeybee (*Apis mellifera* sp) and the *Apis mellifera mellifera* population of the UK, Thesis, University of Leeds.

van Alphen, J.J.M. (2018) Is natural selection alone enough for the evolution of resistance against varroa in the European honeybee?, SICAMM conference, Finland, 2018

Vaha, J. P. and Primmer, C. R. (2006) 'Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci', *Molecular Ecology,* 15(1), 63-72.

Chapter 3

Characterisation of the pollen use and concurrent gut bacteria of Irish honey bees using DNA barcoding

INTRODUCTION

Pollens collected by forager honey bees (*Apis mellifera* L.) contain a colony's primary source of lipids, vitamins, minerals, and proteins including essential amino acids (Herbert, 1992). Their essential amino acids in particular are required in balanced quantities (Degroot 1952, Degroot 1954) since the uptake and utilisation of one amino acid can be dependent on the presence of another (McCaughey *et al.* 1980). Consequently, pollens, either through quantity or nutrient composition, are crucial for the successful development of brood, the hypopharyngeal glands in nurse bees, the immunological response (Herbert *et al.* 1970, Brodschneider and Crailsheim 2010, Graikou *et al.* 2011, Nicolson 2011, Di Pasquale *et al.* 2013, Paoli *et al.* 2014a, Omar *et al.* 2017), the foraging capabilities (Scofield and Mattila 2015) and hence, the health and survival of the colony (Naug 2009, Huang 2012). While the high importance of this food source to colonies is unambiguous, any genetic controls behind pollen choice by bees which may influence the diversity, quantity (Page *et al.* 1995) and nutrient composition of pollen selection are poorly understood.

Individual honey bees are considered to be monolectic, with each bee thought to forage on a singular floral source until the source is exhausted or the bee itself dies. Beekeepers can learn the plants corresponding to the colour of the pollen in the corbicula of returning foragers which, combined with expected floral seasonal availability allows them to better understand how their bees forage the environment surrounding their apiaries (Hodges 1974). This understanding of the monolectic tendency of honey bees is not necessarily correct, however. Inexperienced pollen foragers tend to sample their environment in greater depth and with greater effort than experienced foragers (Pernal and Currie 2001) returning with a mix of pollens. Additionally, as climate change shifts the expected timings of seasons, this can create a disconnect between the expected seasonal flora relative to the actual pollen seen on arriving foragers, making it more difficult for beekeepers to use colour-based indentification. Colonies, as a complete entity, are polylectic, collecting pollen and nectar from a wide range of flowering plants (Hodges, 1974). The angiosperms, flowering plants, are dominantly anemophilous, with entomophily having evolved from this condition (Crawley 1997) with some plants, such as *Castenea* (Chestnuts), which provide large quantities of pollen seen to form a primary source of hive pollen (Aronne *et al.* 2012). Plant species, which may be considered as wind pollinated, can also rely heavily on insect pollination (Tamura and Kudo 2000) and can also form a primary pollen source for colonies. In Ireland, *Salix* (Willows) are a

source of spring pollen for bees, despite being morphologically anemophilous. Nectariferous species which provide both pollen and nectar, feature prominently and entomophilous species make up a final, smaller, proportion (Coffey and Breen 1997).

Considerable variation exists in the nutritional value of pollen (Schmidt 1984, Roulston and Cane 2000, Corby-Harris *et al.* 2018) and the essential nutrients, including amino acids, do not exist in all pollens (Herbert *et al.* 1970). That the honey bee has adapted to forage for pollen diversity where possible (Schmidt 1984) indicates that, apart from pollen quantity, the diversity and quality of pollens foraged by a colony are likely to be important factors in determining fitness (Sedivy *et al.* 2011). Since a diet high in amino acids is required in the early stages of an adult bee's life (Paoli *et al.* 2014b), annual temporal changes in the age profile of a colony may also influence the quantity and also the quality of its pollen requirements at any given time. A detailed understanding of the changes in pollen collection over the foraging season will allow us to better understand the interdependance between honey bees and their environment. The factors influencing pollen selection at a colony level are only partially understood (Aronne *et al.* 2012). Whilst, as previously mentioned, pollen quantity is a key to a strong colony and features highly in beekeepers' colony assessments, careful consideration needs to be given to pollen quality, particularly in the pollens collected in smaller amounts, as these may be for the acquisition of some essential nutrients which may be rarer in the foraging landscape. Although it appears that individual foragers do not necessarily select pollen based on nutritional quality (Pernal and Currie 2001), genetically controlled foraging behaviours (Hunt *et al.* 1995, Page *et al.* 1995, Ben-Shahar *et al.* 2003) may form the basis for differences between the health of colonies through variations in nutritional stress.

An aim of this study is to investigate potential differences between colonies that could factor into their relative health and survival in terms of the pollens they forage. The foraging behaviour of honey bees is driven by a wide range of factors (Abou-Shaara 2014) including the brood state where high brood numbers combined with empty pollen cells prompt the recruitment of additional pollen foragers. Greater genetic diversity within the colony leads to increased colony health by conferring greater immunity to disease (Tarpy 2003, Shykoff and Schmidhempel 1991, Simone-Finstrom et al. 2016). However other mechanisms, consequential on genetic diversity, which may confer increased colony health are possible. Outside of environmental effects such as a climate which curtails flight distance or floral availability, the diversity of pollens collected by a colony may be a function of the variation in

the genetic predisposition of foragers for pollen collection (Page and Fondrk 1995, Pankiw and Page Jr 2000). Greater genetic diversity within a colony's pollen foragers may return a greater diversity in pollens ensuring all forms of essential amino acids and other nutrients are available for colony health.

Increasing numbers of studies are investigating the relationship between land use and pollen foraging. The type and seasonal availability of pollen sources will vary considerably from country to country and region to region making localised studies of pollen sources of critical importance to the understanding of their influence on colony fitness. In Ireland, although there have been studies of the pollen profiles in honey with a view to determining and verifying the honey's origin (Downey *et al.* 2005), to date there has only been a single extensive study of the pollen and nectar sources used by honey bees in Ireland. This study (Coffey and Breen 1997) identified an impressive 106 pollen types and laid a foundation for our understanding of the seasonal use of pollen between regions and colonies in Ireland. However, due to the limitations of light microscopy only 16 pollens were identified to species level with the remainder identified to genus or family.

The use of light microscopy techniques for pollen identification, especially to species level, requires extensive knowledge of the regional plants under investigation and a reference collection, both of which take a considerable amount of time, even years, to accumulate. Even for experienced operators the tasks of identification and quantification, where constant shifts between magnifications are required to identify specific structures of the pollen grain, are highly time-consuming and this has enticed many to seek a more automated (France *et al.* 2000) and less subjective alternative. The proposed use of molecular techniques for barcoding plant species (Stoeckle 2003) and, in particular, the use of the internal transcribed spacer (ITS) regions of nuclear ribosomal DNA (18S, 5.8S and 26S) is a relatively recent development. The ITS regions can be amplified in one of two sections, ITS1 or ITS2, and focus has recently turned to the use of ITS2 barcoding of pollen in ecological plant studies and melissopalynology (Kress *et al.* 2005, Sickel *et al.* 2015, Richardson *et al.* 2015, Smart *et al.* 2017, Keller *et al.* 2015).

It was expected that the ITS2 barcoding used here would identify a greater proportion of the pollens to a lower taxonomic level than was possible for Coffey and Breen (1997).

The south-east of the country has a climate which is more beneficial to honey bees relative to the western counties as it has a higher annual mean temperature, greater total sunshine hours and lower annual rainfall (Appendix I, Pg. 139). The aim of using apiaries from these

two areas was to discover if regional differences in pollen use existed between the west and southeast of the country using the potential additional taxonomic detail provided by barcoding over light microscopy to provide answers. It was hypothesised that western colonies would collect a lower volume of pollen as a consequence of less favourable foraging conditions but that they would collect a greater diversity since there is less crop farming in the west and potentially a greater need for colonies to intensively forage on smaller individual sources. It was also intended to investigate how pollen collection varied within an apiary in volume, diversity and species utilised.

The pollen a colony consumes, collected through active selection by foragers or as determined by availability (McFrederick *et al.* 2012, Donkersley *et al.* 2018), may influence the composition of their gut bacteria (Jones *et al.* 2018). Firstly, through inoculation of the host bees by phyllosphere bacteria (Vorholt 2012) associated with particular pollens (Corby-Harris *et al.* 2014) and secondly by the need for the gut to sustain particular bacteria to facilitate the digestion of particular pollen compounds (Zheng *et al.* 2019). The taxonomic diversity of bacteria in the bee gut may, therefore, be partially a function of the pollen-based nutrients available to them. The relationship between beneficial gut bacteria and the nutrients necessary to sustain them while optimising an individual's health is actively utilised by promoting the growth of beneficial gut bacteria in human bacteriotherapy treatments for illnesses associated with bacteria imbalances (dysbiosis) (Sartor 2008) and is beginning to be promoted for bees (Crotti *et al.* 2013). Interestingly, outside of the gut, in what might be considered a form of ecosystem-wide symbiosis, the bacterial composition of the phyllosphere may be influenced by foraging bees (Aizenberg-Gershtein *et al.* 2013, Prado *et al.* 2020) indicating the possibility that there is a circular arrangement between soil, plant and pollinator which from the perspective of the pollinator meta-population enables 'offsite storage' of beneficial bacterial species.

Before storage by the colony, collected pollen is mixed with nectar which, during its transit time in the honey bee crop, or foregut, will have been inoculated with *Lactobacillus* bacteria (LAB) (Martinson *et al.* 2012, Moran *et al.* 2012). Glandular enzymes are also added to the nectar. It is hypothesised that fermentation of the pollen by the LAB and *Bifidobacterium* genera enable its suitability as larval food (beebread) and also prevents spoilage by fungal pathogens (Olofsson and Vasquez 2008, Vásquez and Olofsson 2009).

While the crop is not involved in digestion (Crailsheim 1988) it does form a common point of contact between the individual bee, the environment (Hannula *et al.* 2019, McFrederick *et al.* 2012) and its nestmates through oral trophallaxis (mouth-to-mouth food transfer). This food transfer behaviour, common in eusocial insects, provides the horizontal transfer of bacteria throughout a colony, facilitating homogenisation of the gut flora between nestmates. This preparation of the colony's main food source is one of the first steps where the bacterial gut flora of honey bees begins to influence colony fitness. Honey bee nutrition is further assisted by bacteria in the mid and hindguts through the efficient breakdown of the hard outer sporopollenin walls of the pollen, facilitating the efficient digestion of the contents within. That only a few bacterial types are capable of sporopollenin breakdown (Mackenzie *et al.* 2015) indicates this is a highly specialised arrangement. Within colonies divergent diets create differences in gut flora (Kapheim *et al.* 2015) indicating that consumed pollen requires a specific suite of gut bacteria to enable its proper digestion and uptake by the host bee. It follows that negative alteration to the optimum bacterial suite, or dysbiosis (Sartor 2008, Hamdi *et al.* 2011), could reduce the health of individual bees and by extension, the fitness of the colony. The importance of the relationship is evidenced by what is, globally, a highly conserved bacterial diversity in the honey bee gut, forming a mutualist symbiosis (Cox-Foster *et al.* 2007, Martinson *et al.* 2011, Moran *et al.* 2012). Dysbiosis, whether qualitative or quantitative, in this diversity raises the possibility of a reduction in survival (Cox-Foster *et al.* 2007) and reproductive fitness (Gavriel *et al.* 2011) of individuals. Shifts in the gut bacterial profile could be expected with shifts in pollen intake such as when the availability of flowers changes seasonally. By examining concurrent pollen and gut bacteria samples the hypothesis in the current study was that the gut bacteria profile would alter qualitatively and/or quantitatively with comparative shifts in the pollen consumed by the adult bees.

MATERIALS AND METHODS

Hive selection

Eight beekeepers (MN, PD, PJC, JH, SK, JL, NL and PJM), each with extensive beekeeping experience and able to commit to a long-term project, were identified. Each beekeeper facilitated the use of an apiary which consisted of a minimum of ten colonies. Six of the apiaries (Coolmore (MN), Youghal (PD), Dungarvan (PJC), Campile (JH), Fethard (SK) & Hacketstown (JL)) were in the Southeast of Ireland and two (Loughrea (NL) & Labane (PJM)) in the West of Ireland (Figure 1).

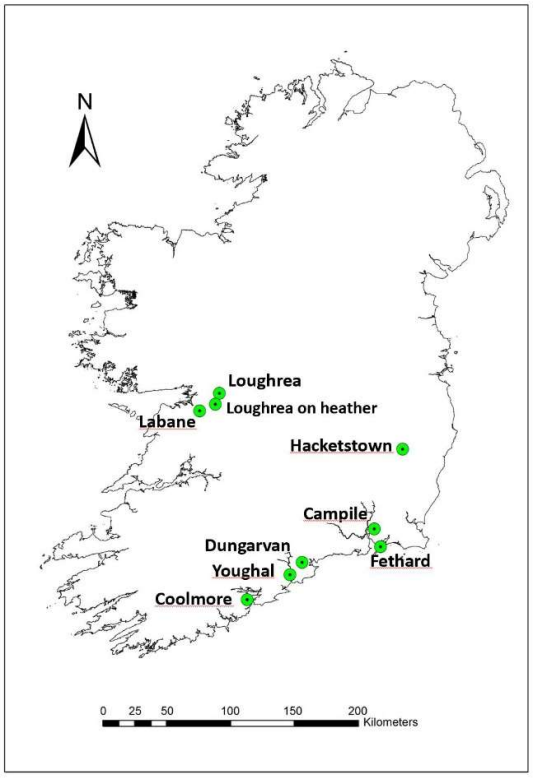


Figure 1: The locations of the eight apiaries used for pollen collection. Altitudes above sea level in metres: Coolmore 3m, Youghal 72m, Dungarvan 26m, Campile 44m, Fethard 10m, Hacketstown 117m, Labane 17m, Loughrea 62m, Loughrea on heather 154m.

All of the apiaries were located in areas of mixed farmland although five had a coastal ecosystem or a national park within a 5km radius. Six of the apiaries were stationary and two, Coolmore and Loughrea, practiced transhumance with a short spring move to *Brassica napus* (Oilseed rape) and a September move to *Calluna vulgaris* (Ling heather) respectively. From each apiary, three hives were chosen at random for pollen sampling (n = 24). Access to the hives was given on the understanding that the beekeepers could continue with their existing management practices and that all efforts would be made to avoid damage or injury to the colonies. This meant that opening the hive during rain or low temperatures was not possible. Seven of the apiaries treated for V*arroa*

destructor using Apiguard® (a thymol based acaricide), oxalic acid or a combination of the two. A single apiary, Coolmore, used no treatment.

Pollen collection

The base section of each hive was replaced with a new base which included an added level to strip pollen from forager's corbiculae (Modification by Donegal bees Ltd.) and an alternate exit/entrance to direct returning foragers into the main hive via the pollen stripper (Figure 2).

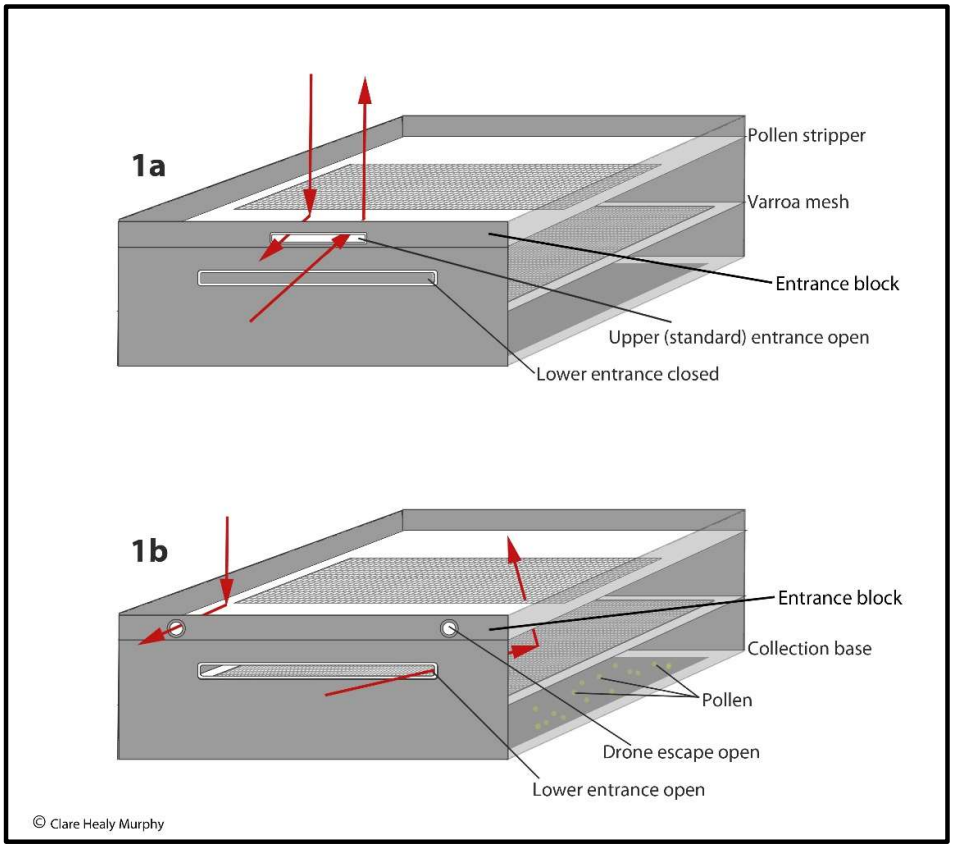


Figure 2: Replacement hive bases used for pollen collection. Bee movement paths are indicated by red arrows.
1a) **Base set to normal operation**: The entrance block is set with a standard open upper entrance and the lower entrance closed. Bees enter into and leave from the brood box above in normal fashion.
1b) **Base set for pollen collection**: The entrance block is turned to closed, leaving only two circular drone escape holes. The lower entrance is opened and returning bees enter into the space under the pollen stripper. To access the main hive returning forager bees must pass through the pollen stripper. Pollen removed from the corbiculae fall through the varroa mesh floor and collect on the lower collection board.

Otherwise, it was a standard base with a varroa mesh and a bottom board. Pollen removed from the corbiculae of returning foragers fell through the varroa mesh and onto the base (collection) board. In normal hive operation the alternate entrance was blocked with foam padding and the bees used a standard entrance in an entrance block. Collection involved

opening the alternate entrance and turning the main entrance block to the closed position, except for two drone escape holes. This arrangement required the majority of returning foragers to pass through the pollen-stripper which removed the fresh pollen from their corbiculae without harming the bee.

To harvest the pollen, the collection board was carefully removed and emptied into a plastic bag. To prevent any sample loss the collection bag was large enough to encompass the entire board.

Between June 2016 and July 2017 inclusive, fresh pollen was collected from returning foragers for three consecutive days approximately every four weeks and stored at -20^{o}C. An assay to randomly collect beebread from frames was abandoned early on (Appendix II, Pg. 142)

Bee collection

Worker bees were collected off a central frame within the brood box. The chosen frame was struck sharply once with the hand to encourage older forager bees to disperse to create a greater density of younger house bees including nurse bees. The frame was then struck a second time over a bucket to collect the bees together wherefrom a sample of between 30 to 40 bees was taken. Bees were taken from the hives approximately every four weeks between June 2016 and July 2017 excluding the winter and early spring (November 2016 to March 2017 inclusive). The samples were cooled immediately upon collection and in transit before being stored at -20°C

Samples were selected from the total collections to represent four seasons (Summer and Autumn 2016, Spring and Summer 2017) and from three apiaries to represent as complete a year as possible within the sampling restrictions.

Pollen preparation and DNA extraction (See Appendix III, Pg145, for the full protocol)

Each pollen sample was transferred to a clean beaker and distilled water added. The mixture was stirred using a magnetic stirrer for 5 minutes and then filtered through a 500µm nylon mesh to remove any hive detritus collected with the pollen.

The filtrate was kept stirred to ensure suspension uniformity while sub-samples were taken from the centre of the mixture using a micropipette. Additional 1mL sub-samples were taken and stored for future use; two to determine pollen density, two for any future microscopy analysis and two for any repeat DNA extractions that may be considered necessary.

For the DNA extraction used in this study a further sub-sample of approximately 50mg of pollen was taken.

For whole DNA extraction, the NucleoSpin® food extraction kit (Machery-Nagel) standard protocol was used, with a prior bead beating step using a SPEX®Sampleprep MiniG® and a mixture of bead sizes to disrupt the variously sized pollen grains.

Bee gut removal and DNA extraction

Working in a laminar flow cabinet, the complete alimentary tracts (guts) were aseptically removed from a pool of five bees (Moran *et al.* 2012) per sample point. To sterilise their exterior surfaces of environmental bacteria each bee was first immersed in a 1% hypochlorine bleach solution and agitated to remove air bubbles trapped on the body surface before being left steep for two minutes. The bee was then washed in three successive baths of distilled water to remove the bleach. To remove the gut the final distal tergite and stinger of each bee was grasped together firmly using forceps and the gut removed in one gentle continuous pulling motion. Where the condition of the sample precluded this method, the entire abdomen was excised and used. The forceps were flame-sterilised between the extractions of each group sample to avoid cross-sample contamination.

Whole DNA was extracted using the DNeasy® Powersoil® kit (Qiagen) (Formerly sold by MO BIO as PowerSoil DNA Isolation Kit) to permit the data to be included in The Earth Microbiome Project (Thompson *et al.* 2017). The pooled tissue was placed in a sterile 2mL PowerBead® microtube and cell lysis agent added per the manufacturer's protocol. In addition to the manufacturer's protocol, to increase the grinding surface area, 100mg of ~100µm acid washed sand was added to each PowerBead® microtube. The tissue was disrupted in a Mini-G bead beater (Spexprep Ltd.) for four minutes at 1500 rpm before being centrifuged at low speed (1000 rpm) for two minutes to pellet the grinding agents. 400-500µL of supernatant was pipetted into a sterile 1.5mL microtube and DNA extracted per the remainder of the manufacturer's protocol.

PCR and sequencing

Whole DNA was sent to collaborators at the Centre for Computational and Theoretical Biology (CCTB) in Universität Würzburg, Germany for next generation sequencing (NGS) using the

Illumina Miseq platform. Duel-index PCR marker amplifications of the internal transcribed spacer part 2 (ITS2) for the pollen (Sickel *et al.* 2015) and 16s rDNA Section V4 for the gut microbiome (Kozich *et al.* 2013) were used. The respective primers used were, for ITS2, primers, S2F - ATGCGATACTTGGTGTGAAT (Chen *et al*. 2010) and ITS4R - TCCTCCGCTTATTGATATGC (White *et al.* 1990) and for the 16s, primers 16Sf (GTGCCAGCMGCCGCGGTAA) and 16Sr (GGACTACHVGGGTWTCTAAT) (Kozich 2013).

For both markers, the PCRs were carried out in three separate 10µL PCR reactions per sample and recombined after fragment replication to mitigate against PCR bias. The amount of PCR product was normalized and purified with the Invitrogen SequalPrep Plate Normalization Kit (ThermoFisher Scientific). BioAnalyzer 2200 (Agilent) with High Sensitivity DNA Chips was used for verification of fragment lengths for both libraries. The final pools were loaded into 500 cycle Reagent V2 Illumina Miseq cartridges with 5% PhiXv3 control library.

Sequence read quality control, OTU production and taxonomic assignment.

Reads were demultiplexed using Qiime 1 (Caporaso *et al.* 2010) before being forward/reverse merged, and filtered for quality and length (>Q20, and >250bp respectively) using USEARCH v9 (Edgar 2010). For OTU clustering, reads were dereplicated and singletons removed. The resulting reads were de novo clustered and denoised with USEARCH v9 at ≥97% similarity to produce final OTUs. For 16s, 2,045 chimeric sequences were removed at this stage.

For ITS2, The National Center for Biotechnology Information (NCBI) online database was queried using the search parameters; ITS2, internal transcribed spacer 2 and ITS 2, and with sequence lengths between 100 and 2000 base pairs. This yielded 4,249 ITS2 fasta sequences. The OTUs were mapped against this database and ≥97% sequence similarity used to identify species.

For 16s, taxonomy was assigned using the RDP classifier (Wang *et al.* 2007) and the RDP set v16 with default settings. To create a community table, reads were mapped to OTUs again using USEARCH v9 and 97% identity threshold.

From the final OTU community tables Qiime 1 was used to calculate the relative abundance values of the taxa found in each hive.

Pollen relative abundance analysis

Pollen abundance data from Qiime 1 was further categorised into five frequency ranges, four in accordance with Louveaux et al (1978) and an additional category for values below 1% as some studies consider these pollens as incidental collection:

>45% - Predominant pollen

>15 to 45% - Secondary pollen

>3 to 15% - Important minor pollen

1% to 3% - Minor pollen

<1% - Incidental pollen

Pollen Alpha diversity analysis

Using the open source software, R v.3.1.2 (R Core team 2018) , alpha diversity values for Observed (species richness), Shannon and Simpson indices were calculated using *phyloseq* v.1.6.1 (McMurdie and Holmes 2013) and plotted using *ggplot2* (Wickham 2016).

RESULTS

Sampling

Two hundred and thirteen samples of both pollen and worker bees were taken from the 24 hives between June 2016 and July 2017. Of the pollen, three samples were inadvertently destroyed and the sample date on a further three went unrecorded. Reduced sampling coverage at Loughrea apiary meant that the three colonies there could not be used for further analysis in this study. Additionally, one hive in Campile and two hives in Hacketstown died. This left 18 colonies with sufficient samples of bees and pollen from which to obtain data (Tables 1a & b).

Table 1a) Temporal distribution of the pollen samples chosen (n=96) from the three hives at seven apiaries. Loughrea was insufficiently sampled to form part of the final data. Aug-16 = Summer 2016 (n=18); Sep-16=Autumn 2016 (n=16); Apr-17 and M1, M2 & M3 of May-17 = Spring 2017 (n=18); Jul-17=Summer 2017 (n=17)

	Coolmore			Youghal			Dungarvn			Campile			Fethard			Hacketstn	Labane			*Total*
Jul-16							C2		C8											2
Aug-16	H4	H5	H14	PD3	PD6	PD9	C2	C3	C8	JH6	JH17	SK42	SK85	SK212		LG2	M1	M2	M3	18
Early Sept-16							C2	C3	C8											3
Sep-16	H4	H5		PD3	PD6	PD9	C2	C3	C8	JH6	JH17	SK42		SK212		LG2	M1	M2	M3	16
Oct-16							C2	C3	C8							LG2				4
Dec-16							C2	C3	C8											3
Mar-17							C2	C3	C8											3
Early Apr-17																LG2				1
Apr-17	H4	H5	H14	PD3	PD6	PD9	C2	C3	C8	JH6	JH17	SK42	SK85	SK212		LG2				15
May-17	H4	H5	H14				C2	C3	C8								M1	M2	M3	9
Jun-17							C2	C3	C8		JH17					LG2				5
Jul-17	H4	H5	H14	PD3	PD6	PD9		C3	C8	JH6	JH17	SK42	SK85	SK212		LG2	M1	M2	M3	17
Year total	5	5	4	4	4	4	10	10	11	4	5	4	3	4		7	4	4	4	**96**

Table 1b) Location and timing of all bee samples (n=96) taken from the colonies (n=18) in their respective apiaries (n=7). The four seasons are bounded by horizontal lines: Aug-16 (Summer 2016), Sep-16 (Autumn 2016), Apr-17 along with May-17 for Coolmore and Labane (Spring 2017) and Jul-17 (Summer 2017). Apiaries with sufficient samples for an annual analysis are shaded

	Campile		Youghal			Dungarvn			Fethard			Coolmore			Hacketstn	Labane			*Total*
Jun-16	JH17	JH6													LG2				3
Jul-16		JH6				C2	C3	C8											4
Aug-16	JH17	JH6	PD3	PD6	PD9	C2	C3	C8	SK212	SK42	SK85	H14	H4	H5	LG2	M1	M2	M3	18
Early Sep-16						C2		C8											2
Sep-16	JH17	JH6	PD3	PD6	PD9	C2	C3	C8	SK212	SK42	SK85	H14	H4	H5	LG2	M1	M2	M3	18
Oct-16	JH17	JH6													LG2				3
Apr-17	JH17	JH6	PD3	PD6	PD9	C2	C3	C8	SK212	SK42	SK85				LG2				12
May-17	JH17	JH6				C2	C3	C8				H14	H4	H5	LG2	M1	M2	M3	12
Jun-17	JH17	JH6				C2	C3	C8							LG2				6
Jul-17	JH17	JH6	PD3	PD6	PD9	C2	C3	C8	SK212	SK42	SK85	H14	H4	H5	LG2	M1	M2	M3	18
Year total	8	9	4	4	4	8	7	8	4	4	4	4	4	4	8	4	4	4	96

As a consequence of inclement weather conditions which prevented the opening of hives to sample bees or where no pollen had been collected by the colony, 14 colony or sample times

from each 96-sample assay group were not replicated in the other, producing an 85.4% overlap of sample timepoints (Tables 1a & b).

For pollen analysis, between 16 and 18 hives from seven apiaries were sufficiently sampled at single time points representing Summer 2016 (August), Autumn 2016 (late September), Spring 2017 (late April and May) and Summer 2017 (July). Additional samples to cover a full year from Dungarvan were chosen (Table 1a). For the microbiome, 18 hives from seven apiaries were sufficiently sampled at each of the same seasonal timepoints (Table 1b).

Pollen sample weights

The number of samples taken per apiary ranged from 11 to 31 (Table 1) and yielded a total of 5271.07g of pollen (n=96). Individual hive pollen totals ranged from 164.22g at PJM2 in Labane to 431.82g at PJC3 in Dungarvan, although the latter was sampled ten times rather than the four times in Labane. Mean pollen weights for individual hives had a range of 27.96g (PJC8) to 104.99g (PD9). The hive mean collected weight was 60.10g and the mean weight range per apiary was 61.22g ranging from 34.39g (Dungarvan) to 82.72g (Youghal) (Table2).

Table 2: Weights in grammes of pollen samples removed from the 18 hives. All samples (n=96) and those to represent the Four seasons (n=69) are shown separately.

		All samples				Four seasons						
Apiary	**Hive**	**No. samples**	**Total**	**Mean**	**Apiary mean**	**Summer 2016**	**Autumn 2016**	**Spring 2017**	**Summer 2017**	**Total**	**Mean**	**Apairy mean**
Coolmore	MNH4	5	362.50	72.50		12.42	28.67	60.90	128.17	230.16	57.54	
	MNH5	5	185.18	37.04		12.81	19.54	34.19	22.58	89.12	22.28	
	MNH14	4	164.81	41.20	**50.25**	14.81	0.00	16.36	73.24	104.41	26.10	**35.31**
Youghal	PD3	4	378.10	94.53		54.50	12.44	96.86	214.30	378.10	94.53	
	PD6	4	194.62	48.66		48.62	48.15	62.81	35.04	194.62	48.66	
	PD9	4	419.97	104.99	**82.72**	202.45	73.36	85.99	58.17	419.97	104.99	**82.72**
Dungarvan	PJC2	10	320.44	32.04		6.40	49.01	87.81	0.00	143.22	35.81	
	PJC3	10	431.74	43.17		15.76	59.20	119.29	110.75	305.00	76.25	
	PJC8	11	307.54	27.96	**34.39**	1.03	72.94	22.86	23.18	120.01	30.00	**47.35**
Campile	JH6	4	217.45	54.36		29.21	69.62	62.35	56.27	217.45	54.36	
	JH17	5	256.59	51.32	**52.84**	18.01	79.04	74.82	6.34	178.21	44.55	**49.46**
Fethard	SK42	4	208.03	52.01		41.79	32.76	52.62	80.86	208.03	52.01	
	SK85	3	271.32	90.44		74.61	0.00	122.79	73.92	271.32	67.83	
	SK212	4	296.77	74.19	**72.21**	10.17	35.49	130.92	120.19	296.77	74.19	**64.68**
Hacketstown	JLLG2	7	528.65	75.52	**75.52**	60.53	6.64	148.50	113.45	329.12	82.28	**82.28**
Labane	PJM1	4	238.34	59.59		-0.18	124.77	54.62	59.13	238.34	59.59	
	PJM2	4	164.22	41.06		-0.87	61.98	17.90	85.21	164.22	41.06	
	PJM3	4	324.80	81.20	**60.61**	0.34	148.50	29.92	146.04	324.80	81.20	**60.61**
	Total	**96**	**5271.07**			**602.41**	**922.11**	**1281.51**	**1406.84**	**4212.87**		
	Mean		***292.84***	***60.10***	***61.22***	***33.47***	***51.23***	***71.20***	***78.16***	***234.05***	***58.51***	***60.34***

At the four seasonal sample points pollen samples (n=69) totalling 4212.87g were taken from the 18 hives (Table 2). There were three occasions when samples could not be taken from individual hives, MNH14 and SK85 in Autumn 2016 and PJC2 in Summer 2017. There was a considerable difference between the largest and smallest amounts of pollen collected, occurring between the two summer collections. The largest total weight of pollen collected

from all colonies was 1406.84g in summer 2017, which was 2.3 times the lowest, 602.41g in summer 2016. The mean weight of pollen collected per apiary across the four sample times was 60.34g. Of the five apiaries where collections were made from three hives (i.e.: excluding Hacketstown and Campile). Coolmore had the lowest amount (35.31g) while Youghal had the largest (82.72g).

Sequencing results

For the ITS2 (Table 3) there were 2,264,494 each of the R1 and R2 reads produced by the Illumina platform with a mean of 23,588 reads per sample and a range from 295 to 69,209. The R2 reads and R1/R2 combined produced considerably lower throughput than R1 alone due to a high degree of errors in the R2 reads therefore only R1 reads were used. After the quality control steps, 1,010,675 R1 reads were obtained with a mean of 10,528 and a range from 54 to 28,991. Sixteen samples fell below the minimum number of reads (1,000) usually required to produce meaningful data. Clustering produced 820,274 distinct sequences ranging from 23 for sample JH17.6.7 to 27,211 for sample C8.10.6. 80.50% of the sequences were successfully mapped against the reference database to identify 68 separate species (Figure 3).

For 16s a total of 4,314,380 paired-end reads were produced. Quality control and length filtering yielded 2,013,001 reads (Appendix VIII, Pg. 157). Dereplication and singleton removal reduced this to 1,677,331 reads ranging from 5,186 to 47,298 (mean of 17,472 reads). 2,045 chimeric sequences were also removed at this stage. Read clustering at ≥97% similarity produced 149 OTUs ranging from 23 in Dungarvan's hive C2 in early September 2016 to 58 in Campile's JH6 in April 2017 (mean OTU count/hive = 35.54, *SD 6.95*).

Pollen relative abundance distribution

The majority of the pollen species identified in each hive sample were categorised as incidental (<1%) with only one or two species constituting the predominant pollen (>45%). Between one to three species constituted the secondary pollens. This was the typical spread of the relative abundances resulting in 1,015 of the 1,496 identifications being classed as incidental, 74 as predominant and 124 as secondary. The species which constituted each category varied between samples.

Table 3 – Sample reads, quality-controlled reads and distinct sequences (arranged by max. to min. post-QC reads) achieved after read quality control using a Phred Q value of 25 (99.7%) and a maximum expected error of 1. The sample codes eg: H4.8.6 contain the Hive name (H4), Month (August) and Year (201**6**). These data are also arranged by sample number in Appendix IV, Pg147. 'E' before a month such as E9 means the sample was taken early in that month, in this case, early September (month 9).

Sample	R1 & R2 Reads	Post-quality control R1 reads	Distinct sequences	Number of Pollen types	Predominant >45%	Secondary >15% to 45%	Imp. Minor >3% to 15%	Minor 1% to 3%	Incidental <1%
H4.8.6	50455	28991	14312	26	1	1		1	23
M2.9.6	46351	28597	26295	19	1	1			17
C8.10.6	47506	27528	27211	16	1				15
M1.9.6	44389	27300	23962	13	2				11
H14.4.7	47228	26969	23571	18	1	1	1	1	14
M1.8.6	43025	25921	17061	27	1	1	2	1	22
H5.9.6	44139	25910	23713	16	1	1	1		13
C3.10.6	43394	24357	19893	13	1				12
JH6.9.6	44492	23899	18138	15	1	1	1		13
SK42.9.6	38728	23361	19398	12		3	1		8
M2.8.6	46676	22511	17181	27	1	1			25
H4.9.6	35830	22247	20941	14	1				13
M3.7.7	36217	22095	19119	17	1	1		1	14
SK212.8.6	44714	21228	17321	18	1	1	4		12
SK85.8.6	42819	20737	16451	18		3	3	1	11
SK212.9.6	31231	20609	8923	12		2	2	1	7
C2.12.6	49087	18927	12477	25		4	1	2	18
JH17.4.7	27767	17628	15272	15	1	1	1	1	11
SK42.8.6	31823	17590	13735	21		2	5	1	13
C3.7.7	26386	17208	16512	19	2				17
C3.9.6	44658	16719	10063	12	1			1	10
JH17.8.6	27025	16695	8045	18	1	1	1	2	13
PD3.9.6	29885	16278	15814	14	1		1	1	11
M1.7.7	27128	14531	12409	22	1		3	1	17
SK85.4.7	32521	14326	10118	20	1	2		1	16
LG2.4.7	20588	13817	12954	13	1	1	1	1	9
C2.3.7	21502	13725	12740	13	1	2	1		9
C8.12.6	62304	13701	11410	21		2	2	3	14
C2.7.6	26284	13468	9859	22		4	2		16
C3.12.6	33553	13412	12775	25		3	4		18
JH17.9.6	22767	13325	11692	11	1		2		8
C3.3.7	19832	13127	12124	14	1	2		1	10
SK42.7.7	21800	13058	11989	14	1	1	1	2	9
C8.7.6	33423	12439	10568	21	1	1	2		17
LG2.10.6	28836	11709	11507	8	1				7
M3.9.6	21715	11536	11277	11	1	1			9
C2.4.7	19139	11486	9632	19	1	1	2		15
M2.7.7	18577	11447	9602	12	1	1	2	1	7
M1.5.7	22763	11434	8742	19	1		4	2	12
JH6.8.6	21785	11148	6489	18	1	2	1	1	13
PD6.4.7	17917	10995	8041	9		3	2	1	3
H14.7.7	18352	10708	7826	23		3		3	17
C3.6.7	19439	10614	9936	13	1	1		2	9
PD3.7.7	22555	10453	10071	19	1	1		1	16
PD9.9.6	32696	10401	6929	13	1			2	10
PD6.9.6	18013	10395	10267	8	1				7
C3.4.7	20796	10346	8850	15	1	2	1	1	10
LG2.7.7	16361	10069	7916	20	1	1	1	1	16
SK212.4.7	14634	9919	9092	23		3	1		19
M2.5.7	21712	9872	7390	25		1	6	3	15
M3.5.7	20033	9768	7223	27	1	1	4	2	19
H14.8.6	37443	9482	6945	23	1		1		21

C8.3.7	12493	8655	7986	9	1	2			6
JH6.4.7	14278	8421	8040	18	1	1	1	1	14
C8.9.6	15989	8222	7042	7	1				6
C2.10.6	69209	8212	2602	17	1			1	15
C8.4.7	14876	7947	5827	14	1	1	1	1	10
SK42.4.7	12993	7682	7049	18	1	1	3		13
PD9.4.7	13216	7593	7162	18		2	4		12
LG2.E4.7	12764	7392	6298	11	1	1		2	7
H4.4.7	14880	7390	6097	20	1	1	2	1	15
C8.5.7	42834	7203	5978	17		3	3	1	10
JH6.7.7	12093	6079	5159	16		3	1		12
PD6.7.7	22926	5949	5632	19	1	1	1	3	13
H5.7.7	20396	5675	5136	19	1	1	2		15
C3.8.6	32007	5634	3608	23		3	1	3	16
PD9.8.6	21249	5391	1379	10	1		1	2	6
PD3.8.6	26172	4795	2626	10	1		1	1	7
C2.9.6	16311	4705	4643	8	1				7
JH17.7.7	13244	4253	3861	14		3	2		9
PD3.4.7	7052	4206	3971	17	1	1	3	1	11
C2.5.7	8163	4061	3311	14	1	2		2	9
H4.7.7	6415	3846	3541	12	1	2		1	8
C8.7.7	16673	3657	2068	16	1	1		5	9
H14.5.7	10397	3552	2980	17	1	2	2	4	8
PD6.8.6	25627	2939	247	12	1	1	1	4	5
LG2.9.6	7046	2877	2792	8	1				7
C8.8.6	11398	2065	944	22		3	3	1	15
C2.6.7	5982	2065	1786	18	1	1	2	1	13
C8.6.7	3380	1005	877	7	1		2		4
SK85.7.7	31907	974	7641	20	1	2	2	1	14
H4.5.7	10569	899	419	17		3	1	6	7
LG2.6.7	5214	808	522	19	1	1	4	2	11
PD9.7.7	16985	353	95	15		1	9	5	
C3.5.7	2906	340	212	8	1	2		3	2
SK212.7.7	16426	275	197	13	1		8	3	1
C2.E9.6	4946	230	82	9	1	1	2	5	
C2.8.6	19976	229	131	12		2	4	4	2
H5.4.7	7188	229	123	10	1	1	2	3	3
M3.8.6	1685	206	103	9	1	1	4	1	2
H5.8.6	1817	156	49	9		2	5	2	
C3.E9.6	6047	144	70	7	1	1	1	4	
LG2.8.6	295	119	99	8		3	2	3	
C8.E9.6	5945	108	40	10		3	3	4	
JH17.6.7	564	64	23	5	1	1	3		
H5.5.7	3668	54	45	8	1		3	4	
Totals	**2264494**	**1010675**	**820274**						

Taxonomy and relative abundance of all pollen

Sixty-eight plant species were identified from all 96 pollen samples (Figure 3). They consisted of 35 native species, 14 long-term introductions, 16 cultivars and three invasives. Appendix V (Pg.150) contains the full list of the identified species' common names, flowering periods, and status in Ireland.

Five species with relative abundances >5% formed 64.28% of the pollen sampled; *Brassica napus* (Oilseed rape) 25.24%, *Hedera helix* (Ivy) 15.59%, *Rubus idaeus* (Raspberry) 10.40%, *Sinapis arvensis* (Charlock) 8.00% and *Ulex europeaus* (Gorse) 5.05%. A further 11 species formed 25.25% of the total pollen and had relative abundances between 1% and 5% *(Trifolium repens* (White clover), *Filipendula ulmaria* (Meadowsweet), *Acer pseudoplatanus* (Sycamore), *Sambucus nigra* (Elderberry), *Ranunculus bulbosus* (Bulbous buttercup), *Lythrum salicaria* (Purple loosestrife), *Vicia faba* (Broad bean), *Rosa canina* (Dog rose), *Cirsium arvense* (Field thistle), *Brassica oleracea* (Wild cabbage) *and Lotus corniculatus* (Bird's-foot trefoil). These 16 species with relative abundance >1% contained 11 native species, 4 cultivars or their close relatives and 1 long-term introduction. The 52 species forming the remaining 10.47% of pollen each had relative abundances <1%.

Of the eight most abundant species, seven were found in all seven apiaries with varying degrees of abundance (Figure 4) while one, *Sambucus nigra*, was not present in Youghal or Campile. Fourteen species were found exclusively in single apiaries. Five of these, *Populus euphratica* (Euphrates poplar), *Populus trichocarpa* (Californian poplar), *Sambucus palmensis* (introduced Alderflower spp.), *Syringa vulgaris* (Lilac) and *Alnus glutinosa* (Common alder) were only found in Labane in the West of Ireland. Of these only *A. glutinosa* is a native species (see Appendices V or VI). The other four species are introductions although *Syringa vulgaris* is a common plant. Two other apiaries each had four exclusive species: *Acer platanoides* (Norway maple), *Prunus lauroserasus* (Cherry laurel), *Ligustrum vulgare* (Wild privet) and *Erysimum cheiri* (Wallflower) were found only at Dungarvan and *Prunus spinosa* (Blackthorn), *Aesculus camea* (Red horse chestnut), *Ficaria verna* (Lessor celandine) and *Prunus cerasifera* (Cherry plum) only at Fethard. *Tillia americana* (Basswood) was found only at Coolmore.

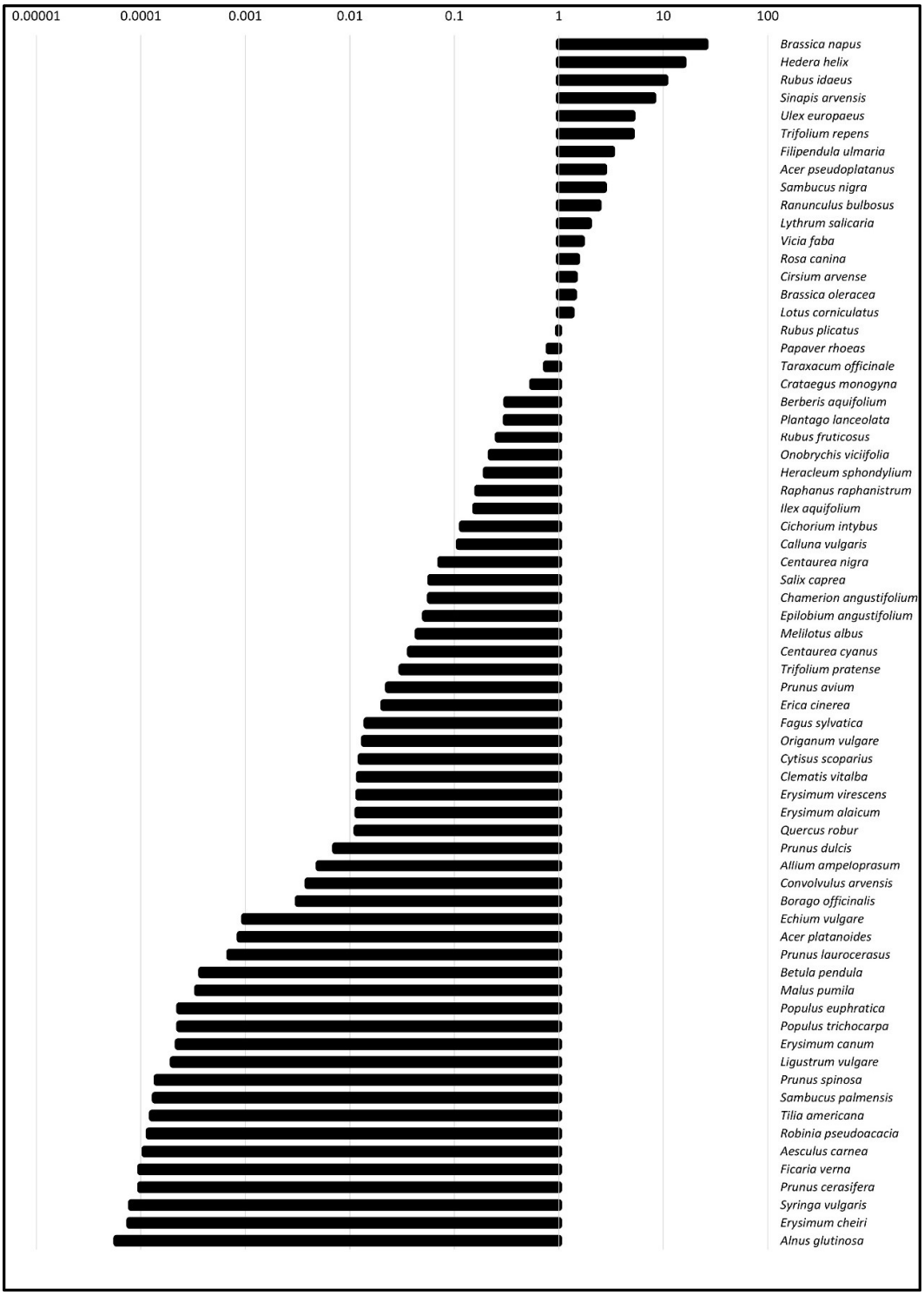


Figure 3: Relative abundance of the 68 pollens identified across all samples (N=96) taken from the seven apiaries. The x-axis values are a logarithmic scale. *Brassica napus* to *Sambucus nigra* form 75.20% of the pollen collected. All species from *Rubus plicatus* down had individual abundance values <1%

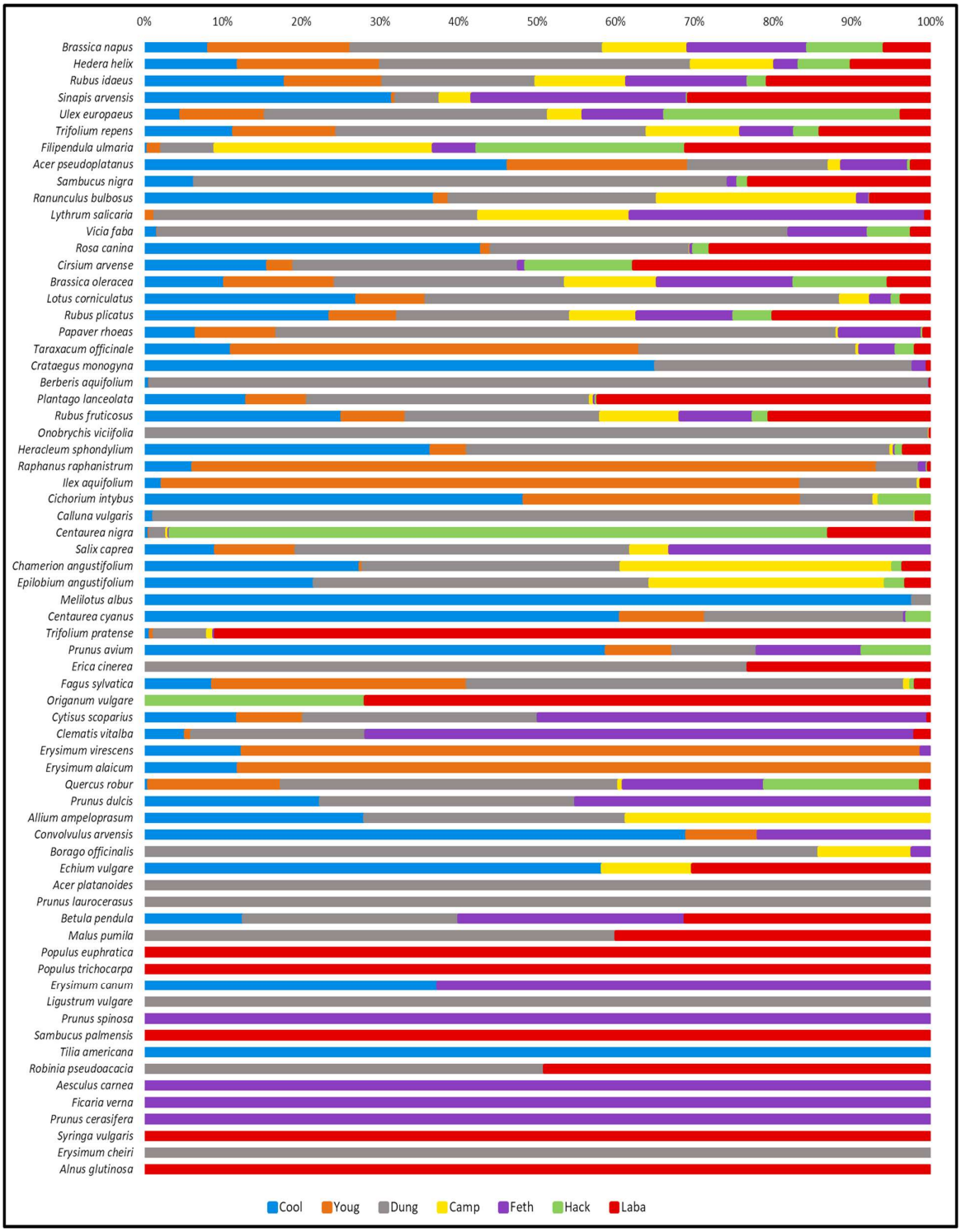


Figure 4: The relative abundance of the species found in the six apiaries. The species are listed in the same order as Figure 3 to aid comparison. Dungarvan had the highest proportion in five of the eight most abundant species. 14 species were found only in a single apiary and are generally confined to the least abundant pollens.

Taxonomic composition of gut microbiome

At phylum level, over three quarters of the microflora identified consisted of a combination of Proteobacteria (51.7%) and Firmicutes (22.1%). Bacteriodetes and Actinobacteria each made up a further 4.7% each. Cyanobacterial sequences, which were all further identified as Chloroplasts, comprised 9.4% (Table 3). A further 6.0% (9 OTUs) of sequences were unclassified.

Table 3: Microbial taxa identified in the 96 samples of pooled bee guts from 21 colonies, collected over a period of one year at seven different locations in Ireland. The number of OTUs from class level that were identified to the lower taxa of order through to species is indicated.

Phylum	% of total	Class	OTUs	Number of each lower taxa identified			
				Orders	Families	Genera	Species
Proteobacteria	51.7	*Alphaproteobacteria*	35	5	11	7	3
		Gammaproteobacteria	31	6	8	13	2
		Betaproteobacteria	10	2	4	3	0
		Deltaproteobacteria	1	1	1	0	0
Firmicutes	22.1	*Bacilli*	31	2	9	9	1
		Clostridia	2	1	2	1	1
Cyanobacteria	9.4	*Chloroplast*	14	2	0	0	0
Bacteroidetes	4.7	*Saprospirae*	2	1	1	2	0
		Cytophagia	1	1	1	1	0
		Flavobacteriia	3	1	1	2	2
		Sphingobacteriia	1	1	1	0	0
Actinobacteria	4.7	*Actinobacteria*	7	2	5	3	0
Verrucomicrobia	0.7	*Pedosphaerae*	1	1	1	0	0
Chlamydiae	0.7	*Chlamydiia*	1	1	0	0	0
		Totals	**140**	**27**	**45**	**41**	**9**

Sixty-nine taxa were identified to genus or above. Most (61) had a relative abundance <1.0%. The 8 taxa constituting 93.91% of the taxonomic abundance and with abundance values >1.0 were: Pasteurelles (Gammaproteobacteria) 34.95%, Lactobacillus (Firmicutes) 29.12%, Acetobacteraceae (Alphabacteraceae) 11.44%, Neisseriaceae (Betaproteobacteria) 9.61%, Bartonellaceae (Alphaproteobacteria) 3.71%, Bifidobacterium (Actinobacteria) 2.09%, Enterobacteriaceae (Gammaproteobacteria) 1.77%, Pseudomonadaceae (Gammaproteobacteria) 1.22% (Appendix IX, Pg. 159).

In individual samples, the relative phylum abundance ranges were, Proteobacteria from 43.4% to 96.5%, Firmicutes from 2.6% to 52.6% and Bacteroidetes and Actinobacteria from 0.0% to 4.6% and 8.3% respectively. Gammaproteobacteria were most abundant class, ranging between 18.1% and 79.2%.

Seasonal abundance

The Summer 2016 and 2017 pollen profiles both indicated higher species richness compared to Autumn 2016 and Spring 2017 in the predominant, secondary, and important minor frequency ranges (Figure 5). There were marked differences between the two summers; 2017 showed a greater use of *Trifolium repens* (White clover) across all apiaries with a mean value per colony of 18.30% compared to 4.57% in 2016, a fact compounded by the 2.3X weight of pollen collected in 2017. Only hive C3 showed considerable use of *T. repens* between both years (24.07% in 2016 and 46.80% in 2017).

Autumn 2016 showed an expected large proportion of *Hedera helix* in all apiaries, along with some *Sinapsis arvensis* and *Brassica napus*. Patterns of pollen use particular to individual apiaries was evident. Both Fethard colonies showed a similar pattern of the three species mentioned above but with the addition of *Brassica oleracea* (Wild cabbage) in roughly equal proportions. The profile of Labane had very consistent proportions of *B. napus* and *S. arvensis* across the three colonies. The single Hacketstown colony had primarily *Ulex europaeus* (Gorse) evident with other species present in very low proportions.

Spring 2017 had greater diversity evident in the pollen profile compared to Autumn. Overall, the cultivar *B. napus* was most abundant along with *U. europaeus* and a reduced signal for *H. helix*. Some within-apiary differences were evident. Hive H5 in Coolmore differed from H4 and H14, showing a high proportion (55.74%) of *Rosa canina* (Dog rose). In Youghal, the profile for hive PD6 showed a greater proportion of *Taraxacum officinale* (Dandelion) compared to PD3 and PD9 (28.64% compared to 1.00% and 6.70% respectively) although it was the only apiary to utilise *T. officinale* in any quantity. The profile for Labane differed considerably from the other apiaries however this apiary was sampled in May rather than April. Superficially the signature spring pollens of the other six colonies (*A. pseudoplatanus, B. napus, B. oleracea* and *U. europaeus*) were present along with the addition of May/June flowering *Sambucus nigra* (Elderberry) along with *R. canina*, *Ranunculous bulbosus* (Bulbous buttercup) and *Plantago lanceolata* (Ribwort plantain) in greater proportions than all the south eastern apiaries except the *R. canina* and *R. bulbosus* proportions in Coolmore.

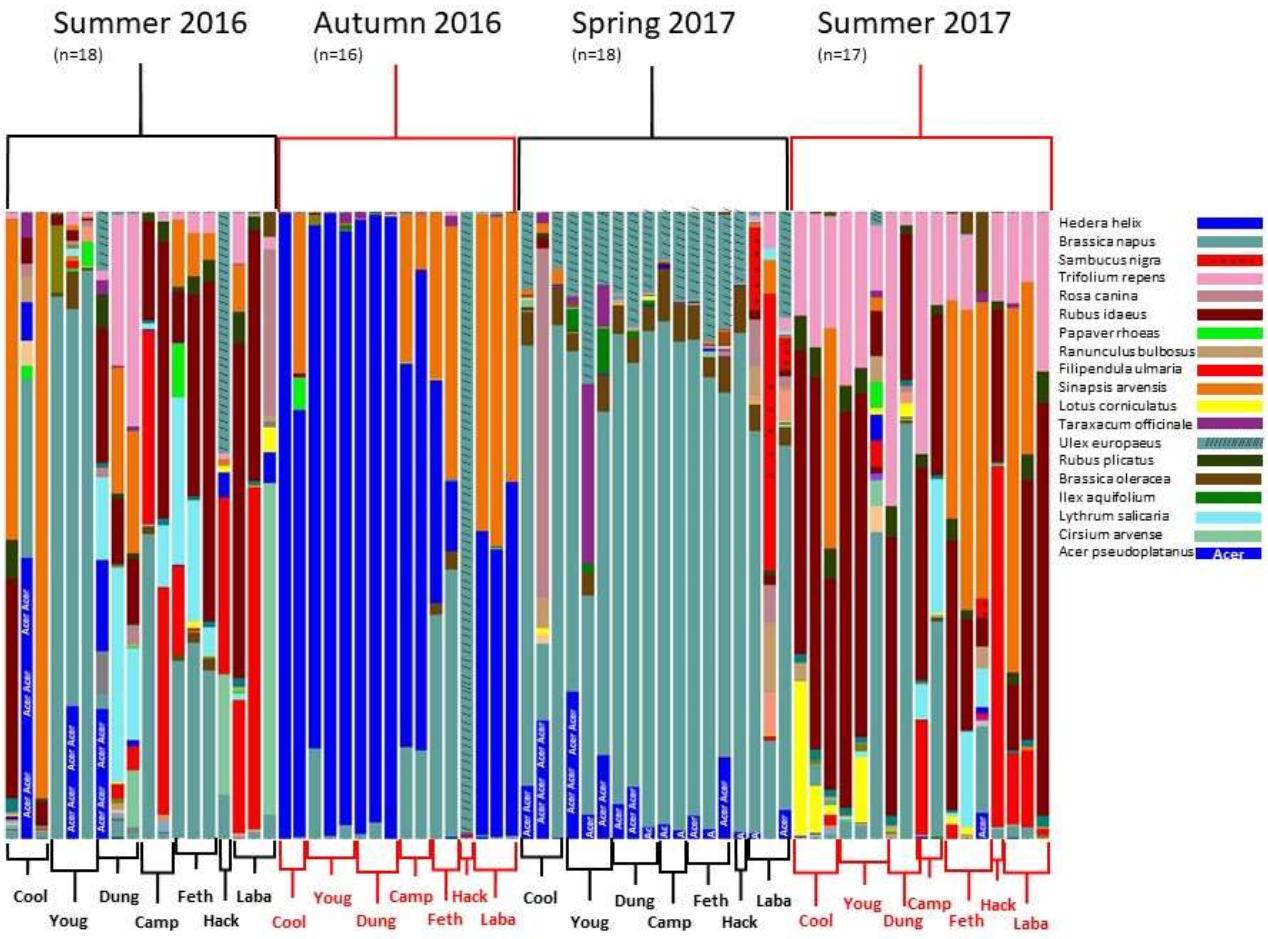


Figure 5: Relative abundance of pollen species for the four seasons between summer 2016 and summer 2017. (See Table 1). Apiaries are: Cool = Coolmore; Youg = Youghal; Dung=Dungarvan; Camp=Campile; Feth=Fethard; Hack=Hacketstown and Laba=Labane. The 19 most abundant pollens are indicated in the legend.

Seasonal gut bacterial composition

Across the four seasons the gut bacteria relative abundance profiles were dominated by the gamma-proteobacterial Pasteuralles, alpha-proteobacterial Acetobacteraceae, the beta-proteobacterial Neisseriaceae and the bacillic *Lactobacillus* (Figure 6). The mean relative abundance of each taxon altered sufficiently between seasons to make seasonal differentiation apparent in the visualisation profiles.

Bifidobacterium genera were present at low frequency in all seasons however there was a marked reduction in their presence in Autumn 2016 relative to the other seasons.

The alpha-proteobacteria Bartonellaceae appeared in varying abundances throughout all seasons. In Spring 2017 however it was only present at minor values except for colony M3 in Labane. Streptophyta values were also noticeably higher at this time.

In Autumn 2016 the Acetobacteraceae signal was generally strongest of the four seasons but with the exceptions of C3 in Dungarvan and JH17 in Campile which both had relatively reduced abundances for this family. In these same colonies Enterobacteriaceae were present in relative abundances of 38% and 25% respectively while the family was not present in any other colonies. Enterobacteriaceae appeared at only one other timepoint, Summer 2017 in Fethard-on-sea (16% in SK212)

The Pseudomonadaceae were present in between two to eight colonies each season, usually at low abundances of between 1% and 5%. In Summer 2017 JH17 in Campile and SK212 in Fethard-on-sea they registered at relatively high abundance values of 27% and 19% respectively. This family appeared in all apiaries at some point in the four seasons except Labane in the West of the country.

Flavobacteriales became more evident in both summers, although not in all colonies at usually at low levels. Of the two summers, they were more evident in 2017 being present in seven out of 17 colonies. They didn’t appear in the Spring 2017 profile.

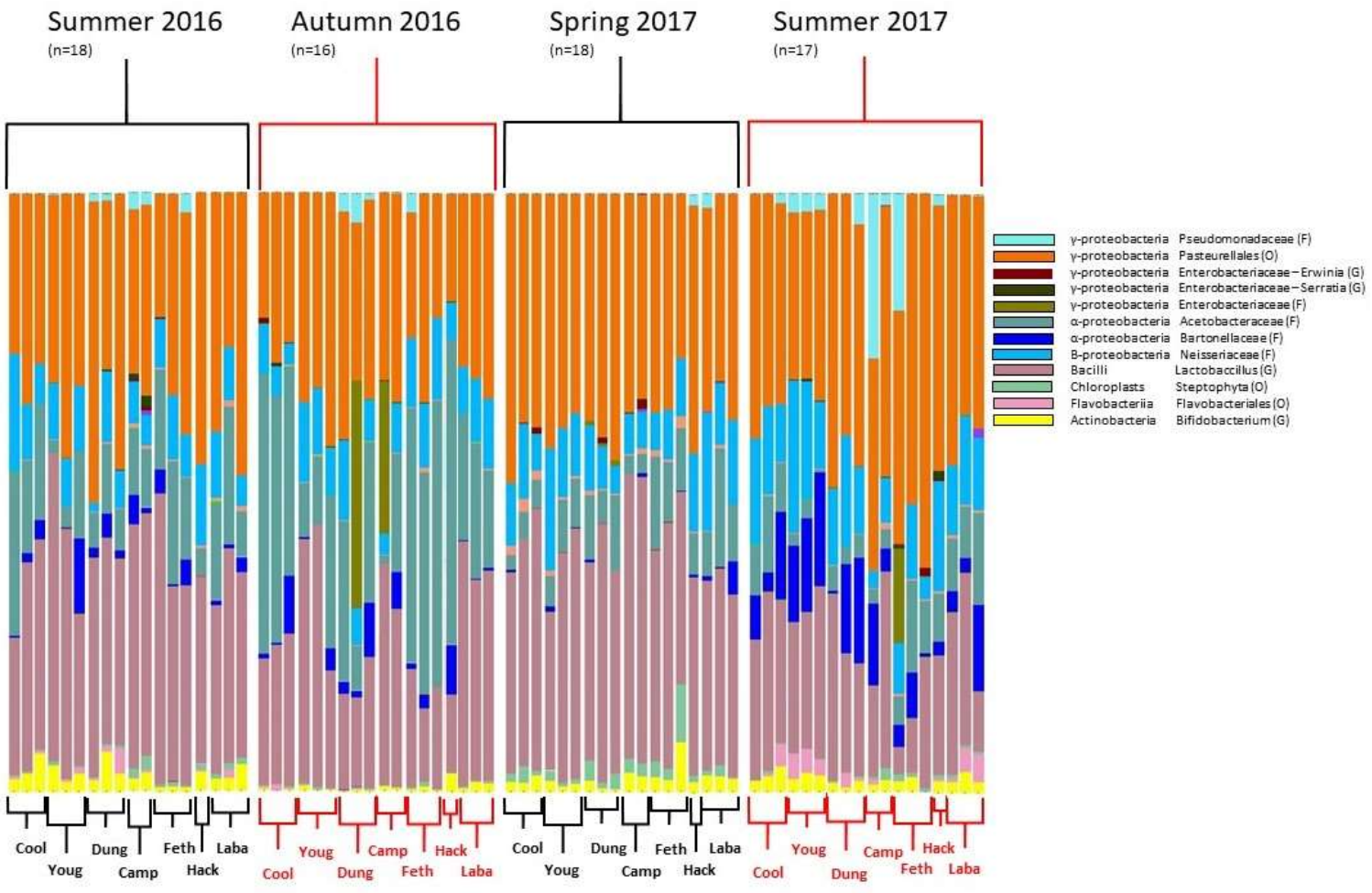


Figure 6 – Microbial composition by relative abundance for the four seasons between Summer 2016 and Summer 2017 as identified to their lowest taxonomic level. O = Order, F = Family, G = Genus. Each multi-coloured column represents a colony within the apiaries. Apiaries are: Cool = Coolmore; Youg = Youghal; Dung=Dungarvan; Camp=Campile; Feth=Fethard; Hack=Hacketstown and Laba=Labane. The legend identifies only the most abundant taxa.

Comparison of pollen within apiaries

Coolmore

For Summer 2016, the Observed diversity showed 27 pollen species for hive H4, 10 for H5 and 20 for H14 (Appendix VII). H4 foraged mainly on *Rubus ideaus* (Raspberry), *Rubus plicatus* (Blackberry) and *Sinapsis arvensis* (Charlock) which was similar to H14 although in different proportions as H14 primarily used *S. arvensis* (Figure 8). H5 differed from the other two colonies, utilising *Acer pseudoplatanus* (Sycamore), *Brassica napus* (Oilseed rape) primarily, combined with small amounts of other species in approximately equal proportions. In summer 2016 H5 had Shannon and Simpson indices of 1.6 and 0.70 respectively (Appendix VII). H4 and H14, which had the larger Observed diversity, had lower Shannon and Simpson diversity indices of 1.0/0.55 and 0.3/0.10 respectively, reflecting the more equal relative abundance in the pollens collected by H5. In Spring 2017, H5 also differed from its apiary-mates, avoiding *B. napus* and collecting a high proportion (55.74%) of *Rosa canina* (dog rose). It was the only hive with a small (6.12%) proportion of *Hedera helix* (Ivy)

Youghal

All three colonies returned comparable proportions of *B. napus* in Summer 2016 however only PD6 had the addition of *A. pseudoplatanus* to any extent (Figure 8). There was a marked difference between the pollen profiles of the two summers which may reflect the later sampling time for Summer 2016, relative to Summer 2017 (Table 1), falling into the flowering season for the winter strain of *B. napus* which has multiple flowering seasons (Wang *et al.* 2011). In Spring 2017 (Figure 8) the three colonies foraged on similar species including the spring strain of *B. napus* which roughly contributed between 34% and 55% of the total pollen per colony. PD6 had a considerably lower Observed diversity (n=10) for spring relative to PD3 (n=17) and PD9 (n=18) however, the Shannon and Simpson indices for all three colonies were similar hovering around the 1.5 and 0.7 marks respectively for all colonies (Figure 7).

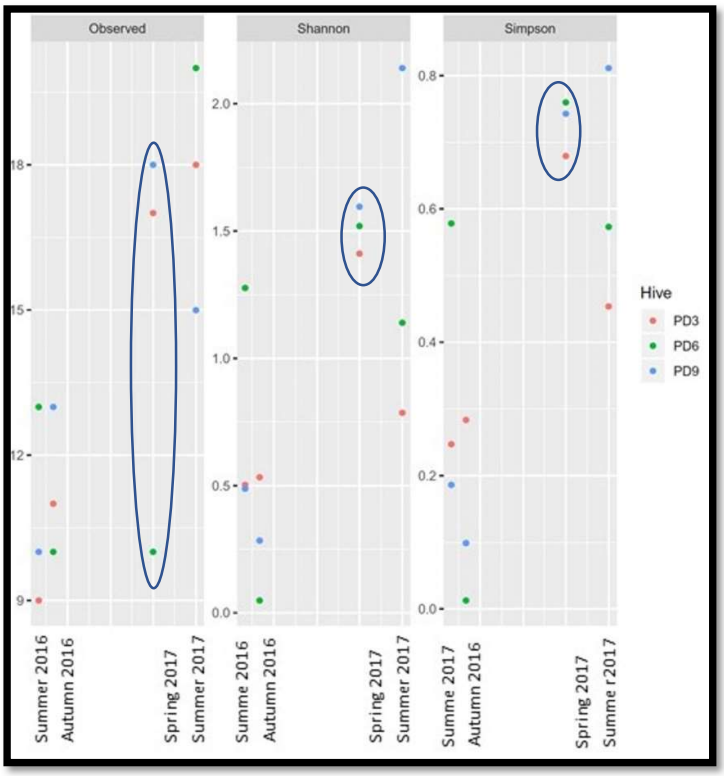


Figure 7: Observed, Shannon and Simpson pollen diversity indices of the Youghal hives plotted for each season. The grouped indices for all three colonies in Spring 2017 are circled in blue showing the compression in index spread when Evenness is accounted for in the Shannon and Simpson values.

Dungarvan

The Autumn 2016 Observed diversity ranged from 4 to 13 however, as *H. helix* formed over 97% of the pollen collected by each of the three colonies, the Shannon and Simpson values for all colonies were around 0.10 for both indices. Spring 2017 consisted of *B. napus* and *U. europaeus* (Figure 8). There was high concordance between the three colonies in the main pollen types collected and their relative abundances. In Summer 2016, colonies C3 and C8 had very similar pollen profiles of *Filipendula ulmaria* (Meadowsweet) *Lythrum salicaria* (Purple-loosestrife), *Rubus idaeus* (Raspberry), *Sinapsis arvensis* (Charlock) and *Trifolium repens* (White clover). While there was some profile overlap with colony C2 it diverged from the other two with *A. pseudoplatanus*, *C. vulgaris*, *H. helix*, and *U. europaeus* forming over 50% of its pollen.

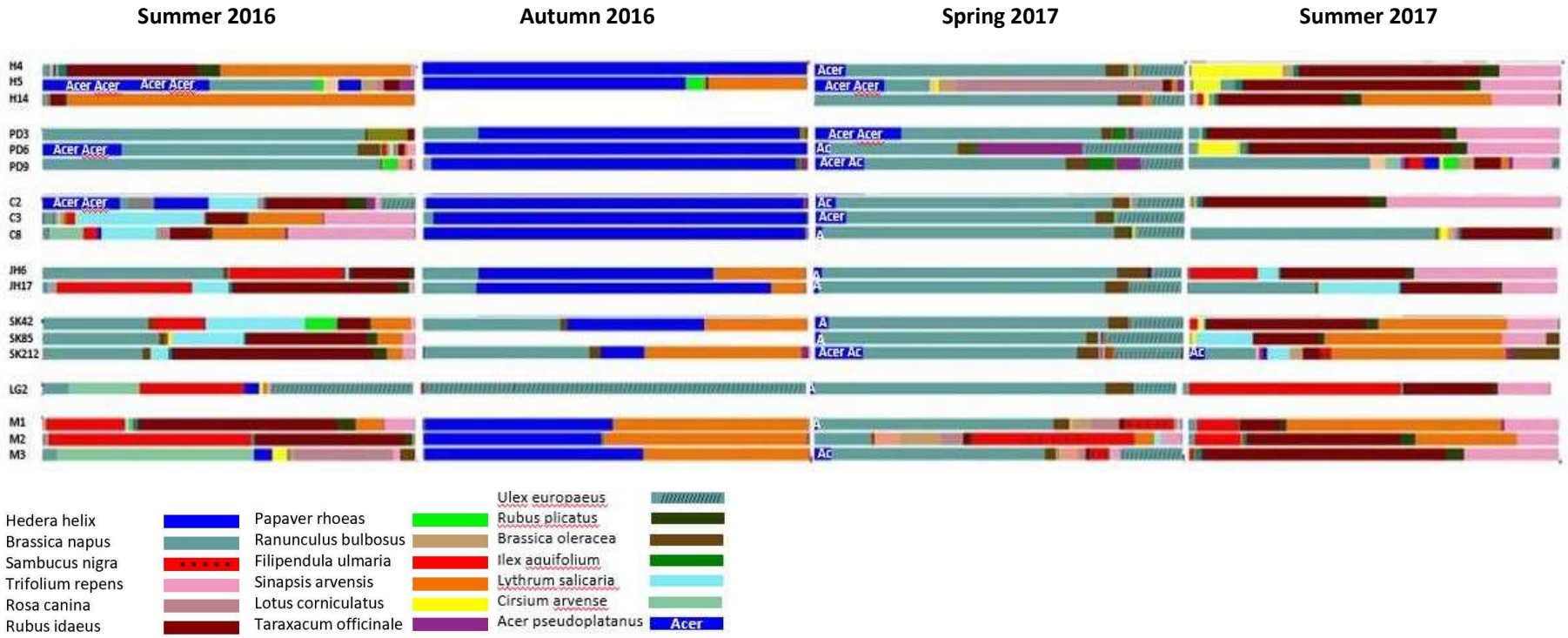


Figure 8 – The relative abundances of pollen species identified for each colony (n=18) of each apiary (n=7) compared across the four seasons sampled. From the top: Coolmore (H4, H5,H14); Youghal (PD3, PD6, PD9); Dungarvan (C2, C3, C8); Campile (JH6, JH17); Hacketstown (LG2); Labane (M1, M2, M3)

Campile

The Autumn 2016 and Spring 2017 profiles were highly concordant between both colonies (Figure 8). In spring the Observed diversity was similar with 17 species for JH6 and 15 species for JH17 and both the Shannon and Simpson indices were almost identical with both colonies at 0.95 and 0.50 (Appendix VII). The pollen profiles for Summer 2016 and 2017 were similar but with important differences. In 2016 both used similar proportions of *F. ulmaria* however, JH6 had a high (48.60%) proportion of *B. napus* whereas JH17 utilised <1%. In summer 2017 only JH6 collected *F. ulmaria* (18.27%) and only JH17 collected *B. napus* (34.43%) in quantity.

Fethard-on-sea

The Autumn 2016 and Spring 2017 profiles (Figure 8) showed a similar use of main pollen types between the SK42 and SK212 colonies. No Autumn 2016 data were available for SK85. In Summer 2016 all three colonies collected varying proportions of *T. repens*, *S. arvensis*, *R. plicatus*, *R.idaeus*, *L. salicaria* and *B. napus* although SK42 differed from the other colonies, collecting significant proportions of *F. ulmaria* (14.17%) and *P. rhoeas* (Common poppy) (8.38%).

Compared to Summer 2016, an initial visual inspection of the Summer 2017 profiles indicated greater divergence between the three colonies. Although the predominant and secondary pollen species used in both years were similar, the proportional use was different. An almost 6-fold mean increase in the proportion of *S. arvensis* from 7.32% to 43.30% and a corresponding 6-fold mean decrease in *B. napus* from 28.85% to 4.76% combined with the addition of the cultivar *Vicia faba* (Broad bean) and *A. pseudoplatanus* created the main differences between the two years.

Hacketstown

The Autumn 2016 pollen was formed almost entirely (99.03%) of *Ulex europaeus* (Gorse), which also respectively formed 38.38% and 11.25% of the Summer 2016 and Spring 2017 collections. The spring pollens were primarily *B. napus* (79.74%) with *A. pseudoplatanus*, *B. oleracea* and *U. europaeus* making up the remainder. Eight species were recorded for Summer 2016 compared to 19 in Summer 2017 which had greater species richness with a Shannon index of 1.1 and Simpson index of 0.59 (Appendix VII). However, as Summer 2017 included

eight species with proportion abundances >0.1% the Shannon and Simpson indices (1.50 and 0.75) both indicated marginally higher diversity for Summer 2016.

Labane

The mean values of the Autumn 2016 pollens for all three colonies were shared equally between *H. helix* and *S. arvensis*. In addition, *U. europaeus*, *Taraxacum officinale* (Dandelion), *Plantago lanceolata* (Ribwort plantain), *B. napus* and *A. pseudoplatanus* were present with individual values <0.5%. In Spring 2017 the three colonies shared a similar palette of species (Figure 8) however the proportional values varied considerably. In Summer 2016 colony M3 differed from M1 and M2. Almost 80% of the pollen for M3 was a combination of *R. canina* and *Cirsium arvense* (Field thistle) whereas the majority (84% to 94%) of the pollens for both M1 and M2 consisted of *F. ulmaria* and *R. idaeus*. In Summer 2017 there was greater concordance in the pollen palette between the three colonies although M3 had very little *F. ulmaria* and no *S. arvensis* with respect to the other two colonies.

The small amounts of additional pollens in Autumn 2016 mentioned above created some variation in the Observed diversity values between all three colonies however their Shannon and Simpson values were almost identical (0.50 and 0.50) reflecting the equality of division between *H. helix* and *S. arvensis* for all colonies, with neither species dominating. In Summer 2017 M1, M2 and M3 had Observed species richness of 23, 12 and 17 respectively. M1 and M2 had similar Shannon indices (1.25 and 1.30 respectively) however, >76% of the M3 pollen species had values ≤1.0% which resulted in a low diversity score (0.40) relative to the other two colonies with the Simpson indices showing a similar division.

Single year case study – Dungarvan

Broad patterns of similarity in species and relative abundances of pollens collected were found between the three colonies in the single Dungarvan apiary. This was particularly so in September and October (Figure 10). Although December, March, and April were also relatively similar, C2 was the only colony that collected a high proportion of *Berberis aquifolium* (Oregon grape) in December. The differences between the colonies at these times were mainly as a consequence of proportional variation in the use of the main pollen types. In March, C2 used mainly *B. napus, B. oleracea, T. officinale and U. europaeus* whiles C3 used the same four species and C8 used three species but not *T. officinale*. The proportional variations, combined with some distinct differences between the colonies such as the use by C2 of the cultivar *Berberis aquifolium* (Oregon grape) in December and *P. rhoeas* in early September and the use of *Lotus corniculatus* (Bird's-foot trefoil) by C3 in early September also, formed the mixed pattern of comparative pollen use. The collection of *B. aquifolium* in December by C2 suggested it could be an early flowering variety since its usual flowering period is February to April. The distinct differences between colonies were usually involved only a single pollen species, however multiple differences between them were also present such as the use of *A. pseudoplatanus* (20.77%), *C. vulgaris* (6.92%) and *H. helix* (14.62%) by C2 in August when none of these were found in the other two colonies.

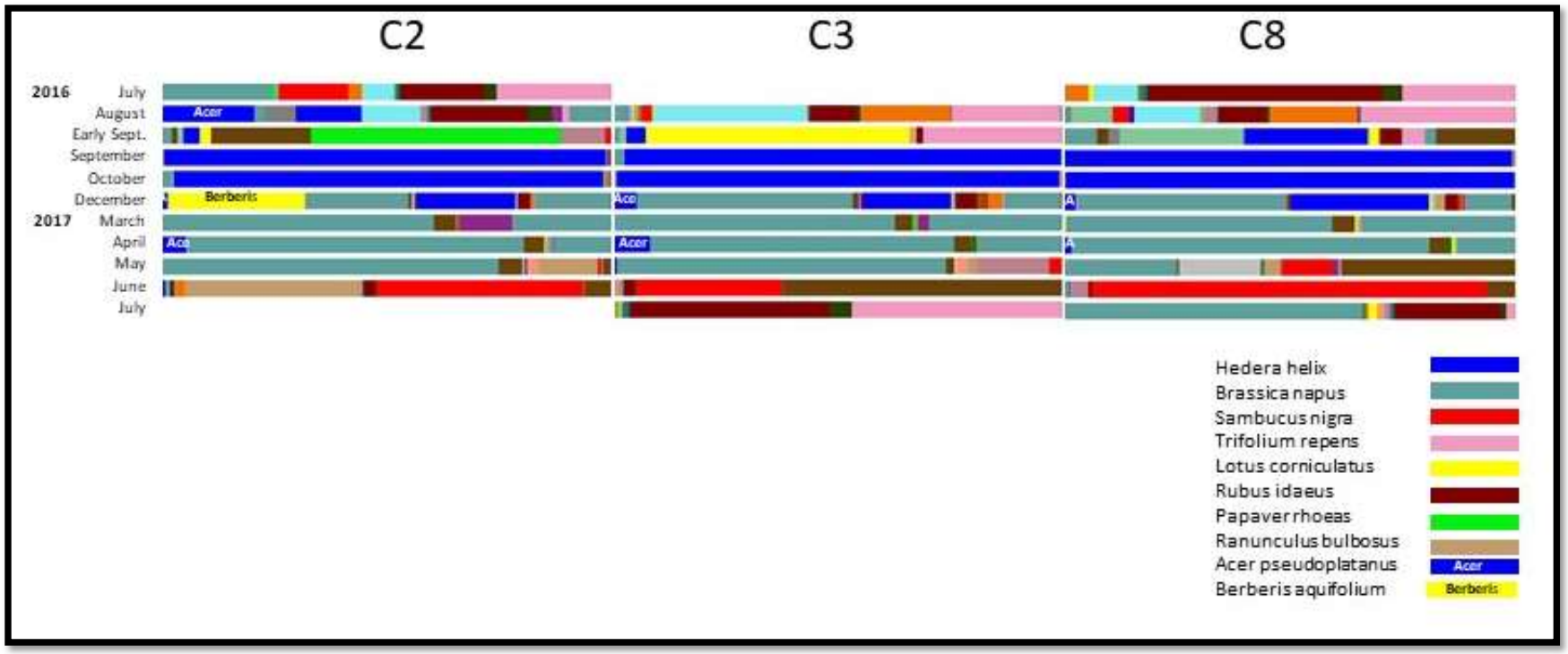


Figure 10: The relative abundance of pollen species collected by the three Dungarvan apiary hives (C2, C3 and C8) from July 2016 to June 2017. The figure is designed to be read from top to bottom and across to allow comparison between the colonies through the year. The nine most abundant pollens are indicated in the legend.

DISCUSSION

Comparison to previous studies in Ireland

The ITS2 molecular barcoding approach to pollen identification used in this study found 820,274 OTUs from which 68 different pollen species from 26 families and 54 genera were identified using the NCBI database. The only previous study of honey bee pollens in Ireland, by Coffey and Breen (1997), identified 92 different pollen types taken from pollen and nectar sources at four sites between 1991 and 1994. These consisted of 24 pollen types identified to Family level, 50 to genus and 16 to species (The remaining two identified were listed as 'types' of an existing taxonomic assignment). Interestingly, there was only a minimal percentage overlap between the two studies. Of the 24 pollen types identified to family level only in 1991-94, only four were identified in this study, a 15% overlap. They were, (synonyms or revised taxonomic assignments are in brackets), Cruciferae (Brassicaceae), Rosacae, Umbelliferae (Apiaceae) and Amaryllidaceae. Fourteen genera (26% overlap) were found in both studies and five species (7% overlap).

It is difficult to explain the reasons behind the low incidence of pollen type overlap which are likely to be the result of a combination of factors. Dungarvan was the only apiary in the current study where every month from May to October was sampled, while the majority of the Coffey and Breen samples were from May to October 1991 at a single site in Templemore, Co. Tipperary. Regional foraging differences between the two locations may account for part of the difference although, against this argument, both locations could be described as "rural, with permanent pasture and small areas of tillage crops", per the Coffey and Breen description of Templemore. The greater diversity of site choice by Coffey and Breen is likely to account for considerable differences. Their study, which included semi-urban and raised bog and one site close to a river, differed from the current study which, although floral diversity variation between sites was present, set out to compare apiaries in the same type of rural landscape and with a single different region for comparison. The light microscopy pollen identification skills of the authors notwithstanding, the improvement in identification particularly to lower taxonomic levels achievable through DNA barcoding (Smart *et al.* 2017) is likely to also be a factor. An overriding advantage of molecular identification over light microscopy has to be the time taken in preparation and identification particularly where DNA extraction can be done in multiples or automated. Once a bioinformatics pipeline is established, such as in 'R' or Qiime, these theoretically do not need significant resetting for

each subsequent study. Conversely, light microscopy studies of pollen, even where micrographs have been prepared and other methods where automation has attempted (France *et al.* 2000, Mitsumoto *et al.* 2009), require new time input for each study. This makes barcoding cost-effective and also more available to the non-specialist (Smart *et al.* 2017) however the extent of identifications is limited by the library(ies) that the OTUs are checked against.

Pollen foraging patterns

Coolmore, Dungarvan, Hacketstown and Labane each had small but significant proportions of *H. helix* collected by single hives in Summer 2016 (respectively H5/6.12%; C8/14.62%; LG2/4.04%; M3/4.90%). Although its collection in August is slightly incongruous with its primary flowering period from September to November (Appendices V or VI), it appears that this pollen source may be important for a larger period of the year than previously thought.

Spring 2017 provides some interesting insight into the flora used by honey bees during colony build up. *Salix* spp. *Corylus avellana* and, later in spring, *Taraxacum officinale* are generally considered the primary sources of colony forage in Ireland's spring however only two colonies, PD6 and PD9 in Youghal, collected significant amounts of *T. officinale* and the other species were not identified in any sample. The high proportions of *Brassica napus* collected (figure 8) may be due to the draw of large amounts of easily available pollen in fields of this crop plant during the early stages of colony build-up but also because its high lipid content makes it extra attractive for bees (Somerville 2005). Where it is practiced, the transhumance of hives to oilseed rape fields for pollination services might cause colonies to temporarily mono-forage this crop although the apiaries examined here, except Coolmore, remained in fixed locations. The three Coolmore colonies were moved to oilseed rape in April 2017 with all of them utilising *B. napus* however colony H5 collected a considerably lower proportion of this pollen than its sister colonies, instead collecting a large proportion of very early flowering *Rosa canina*. This is an interesting development as it involves the use of a plant which is flowering well before its expected earliest period in June, indicating a possible microclimate source such as a walled garden. Additionally, it shows a significant departure in collection pattern of one colony within an apiary despite the universal availability of easy forage. Other untested colonies may of course have followed suit.

It is difficult to distinguish distinct differences between the south eastern and western pollen profiles, although the Spring 2017 results showed the western (Labane) colonies collected *B. napus* they differed from the south-eastern colonies with greater species richness as a consequence of their samples being from May rather than April (Table 1). The difficulty may be due, in part, to the loss of data from the second western apairy (Loughrea) which would have provided a broader picture of western foraging patterns. Another pssible reason is that the Labane apairy was within close foraging distance of the Coole-Garryland Special Protection Area which contains a range of habitats and may have given it increased access to plants that would usually be thinly distributed in western flora. This would have obscured a "standard western" foraging pattern if it exists.

Differences in pollen collection are more discernable between colonies in a single apairy, whether southeastern or western, such as the use of *R. canina* in Coolmore (above). In Dungarvan, only hive C3 showed comparable use of *T. repens* between both summers with 24.07% in 2016 and 46.80% in 2017 (Figure 10). The proportional difference between summers seen in C3 may reflect the one month offset in collection times since 2016 was collected in August and 2017 was collected in July. An additional factor in the difference seen here may be the cyclical use of clovers for fixing nitrogen in arable land. This may be reflected in the greater use of *T. repens* evident across all apiaries in Summer 2017. This is similar to Spring 2017 in Coolmore in that the C3 bees appear to have selected *T. repens* over the alternative *B. napus* foraged by its sister hive C8. Unfortunately, the third hive C3 was not sampled in summer 2017 to allow further comparison within the apiary. A more extensive sampling of all colonies within a single apiary would be needed to elucidate how unusual or not are these differences in collection patterns.

Spring 2017 and both summers had greater species richness evident in their profiles compared to Autumn (Figure 5). This was expected as there are many vernal flowers adapted to take advantage of the lack of canopy shading in spring, while summer is the primary time for flowers to bloom, attracting insect pollinators which function better in warmer weather. In Autumn most plants are fruiting rather than flowering however *Hedra helix* is an exception and becomes the predominant pollen and nectar source across the entire country. This plant, along with *Sinapsis arvensis,* formed the majority pollen in all autumn samples except for colony LG2 at Hacketstown, which only collected *Ulex europaeus*.

Altered flowering phenology

The collection of pollens which were out-of-season according to their generally agreed flowering periods (Appendices V and VI) such as the unexpected appearance of *R. canina* in spring may reflect local conditions or microclimates such as domestic gardens or greenhouses. Plants have peak flowering periods with some individuals opening early and some retaining flowers until well past the peak. Abiotic factors associated with climate change such as increased rainfall and mean daily temperatures can directly affect the flowering period and extend the flowering period 'tails' either side of the peak period to create earlier opening times and later loss of flowers. Additionally the effect that these factors have can combine with others such as habitat removal and artificial soil improvement to reduce floral biodiversity which in turn changes the flowering times of many plants (Wolf *et al.* 2017). It is not clear whether these effects on flowering phenology are sufficient to explain the timings of some of the unexpected pollen appearances in the data recorded here however, irrespective of the causes, it indicates how intimately honey bees investigate their surrounding floral landscape and do not appear to only rely on expected flowerings. Using light microscopy to ground-truth pollen identification was considered if ambiguities about the ITS2-based identification occurred, such as incongruous flowering times, however the training required to attain the skills for the identification of pollens to a taxonomic level sufficient to challenge the molecular data was not considered possible alongside other considerations in the study. Even with the external assistance of an experienced person, the number of pollens likely to be identified to species level would have been considerably fewer than required to provide clear evidence for or against the molecular findings. Multiple sub-samples were stored for microscopy checks to remain a future option.

Incidental pollens

Of the 68 species identified in this study, the relative abundance of 52 of them (75.47%) was <1%, showing that across all apiaries, despite the low diversity of predominant pollens whereby over 72% of the pollen collected came from only nine species (Figure 3), a high diversity of minor and incidental species were collected. The tendency of colonies to collect a few core pollen types points towards oligolectic behaviour at colony level rather than polylectic, however it is worth asking how and why so many "incidental" plant species are

represented. It is possible that this pollen is picked up on a bee's body if they need land on a flower to utilise it as a nectar energy source en route to or from the main pollen forage site (Harano and Sasaki 2015) making it incidental by definition however, the possibility remains that their collection may not be entirely so. Across all 96 samples the mean weight of pollen collected by a single hive over three days was 60.10g (Table 2). If the mean pollen load of a honey bee is ~15mg (Vaissiere and Vinson 1994 ref. Maurizio 1953) then 1% (601mg) represents 40 pollen loads. In some samples there were 18 to 20 pollens with abundances ≤1%, representing a theoretical foraging effort of up to 800 bee-trips from one colony over a three-day period. Amino acids and, more importantly, essential amino acids, minerals and other micro-nutrients are not distributed evenly throughout plant species' pollens (Somerville and Nicol 2006, Roulston and Cane 2000) and it may be that colonies need small amounts of some pollens in order to complete their diet of essential nutrients. Whether these collections involve forager choice, incidental collection with null/positive/negative effect on colony health or wasted effort may be worth further investigation. The evidence is poor for active selection, by individual bees, of pollen based on its quality (Pernal and Currie 2001), however at a colony level, nutritional stress in the hive could conceivably alter the bees' epigenome (Maleszka 2008, Lyko *et al.* 2010, Zhu *et al.* 2017) resulting in an earlier shift of house bees to naïve foragers thereby creating more extensive and diverse foraging for the affected colony which may fill the nutritional need.

Concurrent pollen and colony health data

The current study was restrained to analysis of only a single full year of pollen collections. Idealy the additional data contained in the remaining samples held in frozen repository would add much to our understanding of the pollen use by colonies across the entire year. It was also unfortunately lacking in accurate corresponding health data for each colony such as viral titres or chalkbrood load to compare to the nutrient composition of the incoming pollens, particularly where mono-pollen foraging was evident. Initially, an assay (Appendix II) to sample beebread from within the hive store, combined with a disease assessment, was attempted but abandoned as it was unwieldy and also likely to be detrimental to the health of the colonies to sample too many parameters at each visit. These were not research colonies but formed part of a beekeepers' livelihood. Corresponding colony health data in future studies of pollen use would entail a more controlled sampling environment.

Overview of pollen use

Within apiaries there was broad similarity in the predominant pollen choices between the colonies although single or multiple differences in collected pollen existed and these differences, combined with considerable variation in the choice of minor pollens, makes the picture unclear. The clearest example of this was Youghal, Spring 2017, where a large range of species richness between the colonies was underpinned by a small range of predominant pollens that were even in their proportionality.

Between apiaries, the greatest variation in the predominant and secondary pollen types existed in the summers. Conversely, the Autumn primarily consists of a small range of pollens, even consisting of a single species, such as the singular use of *U. europaeus* by LG2 at Hacketstown, where either the floral landscape may have limited variety, or a particularly attractive nutrition source is available. Similarly, the six south-east apiaries shared a similar palette in Spring with one or two marked exceptions. There was little concordance between the two summers for each apiary, although the timing of the sampling dates is likely to have produced a strong influence on this.

Because Labane was sampled in May rather than April, drawing conclusions for a Spring regional comparison between the West and South-East of the country would be difficult. In both summers there was a large variation between apiaries and between their colonies that it would be difficult to tease out a conclusion without further data, particularly from an additional West Ireland apiary. Notably, autumnal pollen in Labane consisted of the same two pollen types primarily collected in the south-eastern apiaries however the floral landscape would be contracting everywhere at that time of year possibly leading to greater homogeneity in available pollen countrywide.

Future consideration could be given to represent each season's pollens with collections over a wider timescale and with shorter sampling intervals. In this study the single timepoint used was determined to be at the midpoint of each floral season and the intensity of the collections was influenced by the limitations placed on the study through the use of commercial hives where it was necessary to conduct the experiment in such a way that colony health and brood rearing capabilities would not be adversely affected. Additionally, there was wide variation in

the sample weights which would have involved a normalising step to bring each sample to a similar pollen density with others from the same season prior to combining them. As this was an initial use of this molecular method it was decided to keep the analysis as simple as possible and yet create a sufficient picture of honey bee pollen selection in Ireland.

Pollen and gut bacteria interaction

Using the concurrent pollen and gut bacteria samples the hypothesis in this study was that the seasonal gut bacteria profile would alter qualitatively and/or quantitatively with comparative shifts in the pollen consumed by the adult bees. The theory is that as the composition of the food in the bee gut alters, effectively changing niche availability, the bacteria from the phyllosphere inoculating the gut would change to reflect this (McFrederick *et al.* 2012, McFrederick *et al.* 2017, Donkersley *et al.* 2018, Muñoz Colmenero *et al.* 2020). Alternately, the presence or otherwise of particular pollen compounds would act to the benefit or detriment of elements of the incumbent bacterial suite (Zheng *et al.* 2019). This could be an adaptive characteristic, acting as a means to an optimised digestion to changing food availability, particularly where significant seasonal shifts occur. However, any sudden or exaggerated shifts in food source such as could occur in the initial stages of pollen foraging following transhumance or long-range swarming, could initially result in dysbiosis until the colony microbiome adjusts to the new food sources.

Both the pollen and gut microbiome profiles show seasonal variation however that variation is perhaps more subtle in the microbiome as it comes primarily from quantitative shifts in bacterial families rather than from their complete loss or gain in the profile. The pollen sources, flowers, have more distinct changes with the seasons which is likely create a more marked difference. In addition to this, the pollen profiles were distinguishable to species level unlike genus and above as seen in the microbiome and this would allow the profiles to highlight greater taxonomic variation as a consequence of the difference in taxonomic resolution. Additional taxonomic assignments of the bacterial OTUs to species level may reveal variation not seen in this study.

Considered on its own, the composition of the honey bee gut bacteria appears to alter between seasons in the profiles shown here however the reason(s) behind the changes are not evident. There were differences between the two summer profiles such as the increased

presence on Bartonellaceae and Flavobacteriales in 2017 that dilute any argument for a purely seasonal shift due to available flora. Additionally, the pollen use by individual colonies was not identical from one summer to the next although this may reflect the month offset in sampling times between the two summers. It may also be a consequence of the depth to which the landscape was foraged by the bees for each particular year which is dependent on the brood state of the colony at the time of sampling.

Other factors such as pathogen load or disease, differing available flora such as in the cyclical use of clovers or herbicide use (Blot *et al.* 2019) only add to an already complex weave of intrinsic factors. These multiple variables possibly make it extremely difficult if not impossible to accurately determine the existence or otherwise of a direct link between pollen foodstuff and gut bacteria in honey bees except in highly controlled environments. A final point is that the pollen types collected from the corbicula of returning foragers may not be the same as that consumed by the house bees. There is likely to be a time offset between collection and consumption that needs to be accounted for in any future analysis.

References

Aizenberg-Gershtein, Y., Izhaki, I. and Halpern, M. (2013) 'Do honeybees shape the bacterial community composition in floral nectar?', *PLoS One,* 8(7), e67556.

Abou-Shaara, H.F. (2014) The foraging behaviour of honey bees, *Apis mellifera*: A Review. *Veterinarni Medicina*, 59, 1-10

Al-Ghamdi, A. A., Adgaba, N., Tadesse, Y., Getachew, A. and Al-Maktary, A. A. (2017) 'Comparative study on the dynamics and performances of *Apis mellifera jemenitica* and imported hybrid honeybee colonies in southwestern Saudi Arabia', *Saudi Journal of Biological Sciences,* 24(5), 1086-1093.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data [Online]. Available online at www.bioinformatics.babraham.ac.uk/projects/fastqc/

Aronne, G., Giovanetti, M., Guarracino, M. R. and de Micco, V. (2012) 'Foraging rules of flower selection applied by colonies of *Apis mellifera*: ranking and associations of floral sources', *Functional Ecology,* 26(5), 1186-1196.

Ashby, R., Forêt, S., Searle, I. and Maleszka, R. (2016) 'MicroRNAs in honey bee caste determination', *Scientific Reports,* 6, 18794.

Ben-Shahar, Y., Leung, H. T., Pak, W. L., Sokolowski, M. B. and Robinson, G. E. (2003) 'cGMP-dependent changes in phototaxis: A possible role for the foraging gene in honey bee division of labor', *Journal of Experimental Biology,* 206(14), 2507-2515.

Blot, N., Veillat, L., Rouzé, R. and Delatte, H. (2019) 'Glyphosate, but not its metabolite AMPA, alters the honeybee gut microbiota', *PLOS ONE,* 14(4), e0215466.

Brodschneider, R. and Crailsheim, K. (2010) 'Nutrition and health in honey bees', *Apidologie,* 41(3), 278-294.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J. and Knight, R. (2010) 'QIIME allows analysis of high-throughput community sequencing data', *Nature Methods,* 7, 335.

Coffey, M. F. and Breen, J. (1997) 'Seasonal variation in pollen and nectar sources of honey bees in Ireland', *Journal of Apicultural Research,* 36(2), 63-76.

Corby-Harris, V., Maes, P. and Anderson, K. E. (2014) 'The bacterial communities associated with honey bee (*Apis mellifera*) foragers', *PLOS ONE,* 9(4), e95056.

Corby-Harris, V., Snyder, L., Meador, C. and Ayotte, T. (2018) 'Honey bee (*Apis mellifera*) nurses do not consume pollens based on their nutritional quality', *PLoS ONE,* 13(1), e0191050.

Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N. A., Quan, P. L., Briese, T., Hornig, M., Geiser, D. M., Martinson, V., vanEngelsdorp, D., Kalkstein, A. L., Drysdale, A., Hui, J., Zhai, J. H., Cui, L. W., Hutchison, S. K., Simons, J. F., Egholm, M., Pettis, J. S. and Lipkin, W. I. (2007) 'A metagenomic survey of microbes in honey bee colony collapse disorder', *Science,* 318(5848), 283-287.

Crailsheim, K. (1988) 'Regulation of food passage in the intestine of the honeybee (*Apis mellifera* L.)', *Journal of insect physiology.,* 34(2), 85-90.

Crawley, M. J. (1997) *Plant Ecology,* Hoboken, United Kingdom: John Wiley & Sons, Incorporated.

Crotti, E., Sansonno, L., Prosdocimi, E. M., Vacchini, V., Hamdi, C., Cherif, A., Gonella, E., Marzorati, M. and Balloi, A. (2013) 'Microbial symbionts of honeybees: a promising tool to improve honeybee health', *New Biotechnology,* 30(6), 716-722.

Degroot, A. P. (1952) 'Amino acid requirements for growth of the honeybee (*Apis mellifica* L)', *Experientia,* 8(5), 192-193.

Degroot, A. P. (1954) 'Qualitative amino acid requirements of the honeybee', *Acta Physiologica Et Pharmacologica Neerlandica,* 3(3), 433-434.

Di Pasquale, G., Salignon, M., Le Conte, Y., Belzunces, L. P., Decourtye, A., Kretzschmar, A., Suchail, S., Brunet, J. L. and Alaux, C. (2013) 'Influence of pollen nutrition on honey bee health: Do pollen quality and diversity matter?', *Plos One,* 8(8).

Donkersley, P., Rhodes, G., Pickup, R. W., Jones, K. C. and Wilson, K. (2018) 'Bacterial communities associated with honeybee food stores are correlated with land use', *Ecology and Evolution,* 8(10), 4743-4756.

Downey, G., Hussey, K., Kelly, J. D., Walshe, T. F. and Martin, P. G. (2005) 'Preliminary contribution to the characterisation of artisanal honey produced on the island of Ireland by palynological and physico-chemical data', *Food Chemistry,* 91(2), 347-354.

Edgar, R. C. (2010) 'Search and clustering orders of magnitude faster than BLAST', *Bioinformatics,* 26(19), 2460-2461.

France, I., Duller, A. W. G., Duller, G. A. T. and Lamb, H. F. (2000) 'A new approach to automated pollen analysis', *Quaternary Science Reviews,* 19(6), 537-546.

Gavriel, S., Jurkevitch, E., Gazit, Y. and Yuval, B. (2011) 'Bacterially enriched diet improves sexual performance of sterile male Mediterranean fruit flies', *Journal of Applied Entomology,* 135(7), 564-573.

Glastad, K. M., Hunt, B. G., Yi, S. V. and Goodisman, M. A. D. (2011) 'DNA methylation in insects: on the brink of the epigenomic era', *Insect Molecular Biology,* 20(5), 553-565.

Graikou, K., Kapeta, S., Aligiannis, N., Sotiroudis, G., Chondrogianni, N., Gonos, E. and Chinou, I. (2011) 'Chemical analysis of Greek pollen - Antioxidant, antimicrobial and proteasome activation properties', *Chemistry Central Journal,* 5.

Hamdi, C., Balloi, A., Essanaa, J., Crotti, E., Gonella, E., Raddadi, N., Ricci, I., Boudabous, A., Borin, S., Manino, A., Bandi, C., Alma, A., Daffonchio, D. and Cherif, A. (2011) 'Gut microbiome dysbiosis and honeybee health', *Journal of Applied Entomology,* 135(7), 524-533.

Hannula, S. E., Zhu, F., Heinen, R. and Bezemer, T. M. (2019) 'Foliar-feeding insects acquire microbiomes from the soil rather than the host plant', *Nature Communications,* 10(1), 1254.

Harano, K. and Sasaki, M. (2015) 'Adjustment of honey load by honeybee pollen foragers departing from the hive: the effect of pollen load size', *Insectes Sociaux,* 62(4), 497-505.

Hatjina, F., Costa, C., Buechler, R., Uzunov, A., Drazic, M., Filipi, J., Charistos, L., Ruottinen, L., Andonov, S., Meixner, M. D., Bienkowska, M., Dariusz, G., Panasiuk, B., Le Conte, Y., Wilde, J., Berg, S., Bouga, M., Dyrba, W., Kiprijanovska, H., Korpela, S., Kryger, P., Lodesani, M., Pechhacker, H., Petrov, P. and Kezic, N. (2014) 'Population dynamics of European honey bee genotypes under different environmental conditions', *Journal of Apicultural Research,* 53(2), 233-247.

Herbert EW Jr (1992) Honey bee nutrition. In: Graham, JM (ed) The hive and the honey bee. Dadant, Hamilton, Ill, pp 197–233

Herbert, E. W., Bickley, W. E. and Shimanuki, H. (1970) 'Brood-rearing capability of caged honey bees fed dandelion and mixed pollen diets', *Journal of Economic Entomology,* 63(1), 215+.

Hodges, D, (1974) The pollen loads of the honey bee, *International Bee Research Association*, 1974 revision

Huang, Z., (2012) 'Pollen nutrition affects honey bee stress resistance', *Terrestrial Arthropod Reviews,* (5) 175-189

Hunt, G. J., Page, R. E., Fondrk, M. K. and Dullum, C. J. (1995) 'Major quantitative trait loci affecting honey bee foraging behaviour', *Genetics,* 141(4), 1537-1545.

Jones, J. C., Fruciano, C., Hildebrand, F., Al Toufalilia, H., Balfour, N. J., Bork, P., Engel, P., Ratnieks, F. L. W. and Hughes, W. O. H. (2018) 'Gut microbiota composition is associated with environmental landscape in honey bees', *Ecology and evolution.,* 8(1), 441-451.

Keller, A., Danner, N., Grimmer, G., Ankenbrand, M., von der Ohe, K., von der Ohe, W., Rost, S., Hartel, S. and Steffan-Dewenter, I. (2015) 'Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples', *Plant Biology,* 17(2), 558-566.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. and Schloss, P. D. (2013) 'Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform', *Applied and Environmental Microbiology,* 79(17), 5112-5120.

Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A. and Janzen, D. H. (2005) 'Use of DNA barcodes to identify flowering plants', *Proceedings of the National Academy of Sciences of the United States of America,* 102(23), 8369-8374.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C. and Maleszka, R. (2010) 'The honey bee epigenomes: Differential methylation of brain DNA in queens and workers', *PLOS Biology,* 8(11), e1000506.

Mattila, H. R. and Otis, G. W. (2006) 'Influence of pollen diet in spring on development of honey bee (Hymenoptera : Apidae) colonies', *Journal of Economic Entomology,* 99(3), 604-613.

McCaughey, W. F., Gilliam, M. and Standifer, L. N. (1980) 'Amino acids and protein adequacy for honey bees of pollens from desert plants and other floral sources', *Apidologie,* 11(1), 75-86.

Mackenzie, G., Boa, A. N., Diego-Taboada, A., Atkin, S. L. and Sathyapalan, T. (2015) 'Sporopollenin, the least known yet toughest natural biopolymer', *Frontiers in Materials,* 2, 66.

Maleszka, R. (2008) 'Epigenetic integration of environmental and genomic signals in honey bees: the critical interplay of nutritional, brain and reproductive networks', *Epigenetics,* 3(4), 188-192.

Martinson, V. G., Danforth, B. N., Minckley, R. L., Rueppell, O., Tingek, S. and Moran, N. A. (2011) 'A simple and distinctive microbiota associated with honey bees and bumble bees', *Molecular Ecology,* 20(3), 619-628.

Martinson, V. G., Moy, J. and Moran, N. A. (2012) 'Establishment of characteristic gut bacteria during development of the honeybee worker', *Applied and Environmental Microbiology,* 78(8), 2830-2840.

McFrederick, Q. S., Thomas, J. M., Neff, J. L., Vuong, H. Q., Russell, K. A., Hale, A. R. and Mueller, U. G. (2017) 'Flowers and wild Megachilid bees share microbes', *Microbial Ecology,* 73(1), 188-200.

McFrederick, Q. S., Wcislo, W. T., Taylor, D. R., Ishak, H. D., Dowd, S. E. and Mueller, U. G. (2012) 'Environment or kin: whence do bees obtain acidophilic bacteria?', *Molecular Ecology,* 21(7), 1754-1768.

McMurdie, P. J. and Holmes, S. (2013) 'phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data', *PLOS ONE,* 8(4), e61217.

Meixner, M. D., Buechler, R., Costa, C., Francis, R. M., Hatjina, F., Kryger, P., Uzunov, A. and Carreck, N. L. (2014) 'Honey bee genotypes and the environment', *Journal of Apicultural Research,* 53(2), 183-187.

Meixner, M. D., Francis, R. M., Gajda, A., Kryger, P., Andonov, S., Uzunov, A., Topolska, G., Costa, C., Amiri, E., Berg, S., Bienkowska, M., Bouga, M., Büchler, R., Dyrba, W., Gurgulova, K., Hatjina, F., Ivanova, E., Janes, M., Kezic, N., Korpela, S., Conte, Y. L., Panasiuk, B., Pechhacker, H., Tsoktouridis, G., Vaccari, G. and Wilde, J. (2014) 'Occurrence of parasites and pathogens in honey bee colonies used in a European genotype-environment interactions experiment', *Journal of Apicultural Research,* 53(2), 215-229.

Meixner, M. D., Kryger, P. and Costa, C. (2015) 'Effects of genotype, environment, and their interactions on honey bee health in Europe', *Current Opinion in Insect Science,* 10, 177-184.

Mitsumoto, K., Yabusaki, K. and Aoyagi, H. (2009) 'Classification of pollen species using autofluorescence image analysis', *Journal of Bioscience and Bioengineering,* 107(1), 90-94.

Moran, N. A., Hansen, A. K., Powell, J. E. and Sabree, Z. L. (2012) 'Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees', *Plos One,* 7(4).

Muñoz-Colmenero, A. M., Baroja-Careaga, I., Kovačić, M., Filipi, J., Puškadija, Z., Kezić, N., Estonba, A., Ralph, B. and Zarraonaindia, I. (2020) 'Differences in honey bee bacterial diversity and composition in agricultural and pristine environments – a field study', *Apidologie*.

Naug, D. (2009) 'Nutritional stress due to habitat loss may explain recent honeybee colony collapses', *Biological Conservation,* 142(10), 2369-2372.

Nicolson, S. W. (2011) 'Bee food: the chemistry and nutritional value of nectar, pollen and mixtures of the two', *African Zoology,* 46(2), 197-204.

Olofsson, T. C. and Vasquez, A. (2008) 'Detection and identification of a novel lactic acid bacterial flora within the honey stomach of the honeybee *Apis mellifera'*, *Current Microbiology,* 57(4), 356-363.

Omar, E., Abd-Ella, A. A., Khodairy, M. M., Moosbeckhofer, R., Crailsheim, K. and Brodschneider, R. (2017) 'Influence of different pollen diets on the development of hypopharyngeal glands and size of acid gland sacs in caged honey bees (*Apis mellifera*)', *Apidologie,* 48(4), 425-436.

Page, R. E. and Fondrk, M. K. (1995) 'The effects of colony-level selection on the social orgagization of honey bee (*Apis mellifera* L) colonies – Clony level components of pollen hoarding', *Behavioral Ecology and Sociobiology,* 36(2), 135-144.

Page, R. E., Waddington, K. D., Hunt, G. J. and Fondrk, M. K. (1995) 'Genetic determinants of honey bee foraging behaviour', *Animal Behaviour,* 50, 1617-1625.

Pankiw, T. and Page Jr, R. E. (2000) 'Response thresholds to sucrose predict foraging division of labor in honeybees', *Behavioral Ecology and Sociobiology,* 47(4), 265-267.

Paoli, P. P., Donley, D., Stabler, D., Saseendranath, A., Nicolson, S. W., Simpson, S. J. and Wright, G. A. (2014a) 'Nutritional balance of essential amino acids and carbohydrates of the adult worker honeybee depends on age', *Amino Acids,* 46(6), 1449-1458.

Paoli, P. P., Wakeling, L. A., Wright, G. A. and Ford, D. (2014b) 'The dietary proportion of essential amino acids and Sir2 influence lifespan in the honeybee', *Age,* 36(3), 1239-1247.

Parker, R., Melathopoulos, A. P., White, R., Pernal, S. F., Guarna, M. M. and Foster, L. J. (2010) 'Ecological adaptation of diverse honey bee (*Apis mellifera*) populations (adaptation in bees)', *PLoS ONE,* 5(6), e11096.

Pernal, S. F. and Currie, R. W. (2001) 'The influence of pollen quality on foraging behavior in honeybees (*Apis mellifera* L.)', *Behavioral Ecology and Sociobiology,* 51(1), 53-68.

Prado, A., Marolleau, B., Vaissière, B. E., Barret, M. and Torres-Cortes, G. (2020) 'Insect pollination: an ecological process involved in the assembly of the seed microbiota', *Scientific Reports,* 10(1), 3575.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at https://www.R-project.org/.

Richardson, R. T., Lin, C. H., Sponsler, D. B., Quijia, J. O., Goodell, K. and Johnson, R. M. (2015) 'Applications of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem', *Applications in Plant Sciences,* 3(1).

Roulston, T. H. and Cane, J. H. (2000) 'Pollen nutritional content and digestibility for animals', *Plant Systematics and Evolution,* 222(1-4), 187-209.

Sartor, R. B. (2008) 'Therapeutic correction of bacterial dysbiosis discovered by molecular techniques', *Proceedings of the National Academy of Sciences,* 105(43), 16413.

Schmidt, J. O. (1984) 'Feeding preferences of *Apis mellifera* (Hymenoptera, Apidae) – Individual versus mixed pollen species', *Journal of the Kansas Entomological Society,* 57(2), 323-327.

Scofield, H. N. and Mattila, H. R. (2015) 'Honey bee workers that are pollen stressed as larvae become poor foragers and waggle dancers as adults', *PLoS ONE,* 10(4), e0121731.

Sedivy, C., Muller, A. and Dorn, S. (2011) 'Closely related pollen generalist bees differ in their ability to develop on the same pollen diet: evidence for physiological adaptations to digest pollen', *Functional Ecology,* 25(3), 718-725.

Shykoff, J. A. and Schmid-hempel, P. (1991) 'Genetic relatedness and eusociality – parasite mediated selection on the genetic composition of groups', *Behavioral Ecology and Sociobiology,* 28(5), 371-376.

Sickel, W., Ankenbrand, M. J., Grimmer, G., Holzschuh, A., Hartel, S., Lanzen, J., Steffan-Dewenter, I. and Keller, A. (2015) 'Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach', *Bmc Ecology,* 15.

Simone-Finstrom, M., Walz, M. and Tarpy, D. R. (2016) 'Genetic diversity confers colony-level benefits due to individual immunity', *Biology Letters,* 12(3).

Smart, M. D., Cornman, R. S., Iwanowicz, D. D., McDermott-Kubeczko, M., Pettis, J. S., Spivak, M. S. and Otto, C. R. V. (2017) 'A comparison of honey bee-collected pollen from working agricultural lands using light microscopy and ITS metabarcoding', *Environmental Entomology,* 46(1), 38-49.

Somerville, D. (2005) 'Lipid content of honey bee-collected pollen from south-east Australia', *Australian Journal of Experimental Agriculture - AUST J EXP AGR,* 45.

Somerville, D. and Nicol, H. (2006) 'Crude protein and amino acid composition of honey bee-collected pollen pellets from south-east Australia and a note on laboratory disparity', *Australian Journal of Experimental Agriculture - AUST J EXP AGR,* 46.

Stoeckle, M. (2003) 'Taxonomy, DNA, and the bar code of life', *Bioscience,* 53(9), 796-797.

Tamura, S. and Kudo, G. (2000) 'Wind pollination and insect pollination of two temperate willow species, *Salix miyabeana* and *Salix sachalinensis*', *Plant Ecology,* 147(2), 185-192.

Tarpy, D. R. (2003) 'Genetic diversity within honeybee colonies prevents severe infections and promotes colony growth', *Proceedings of the Royal Society B-Biological Sciences,* 270(1510), 99-103.

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J. T., Mirarab, S., Zech Xu, Z., Jiang, L., Haroon, M. F., Kanbar, J., Zhu, Q., Jin Song, S., Kosciolek, T., Bokulich, N. A., Lefler, J., Brislawn, C. J., Humphrey, G., Owens, S. M., Hampton-Marcell, J., Berg-Lyons, D., McKenzie, V., Fierer, N., Fuhrman, J. A., Clauset, A., Stevens, R. L., Shade, A., Pollard, K. S., Goodwin, K. D., Jansson, J. K., Gilbert, J. A., Knight, R., Rivera, J. L. A., Al-Moosawi, L., Alverdy, J., Amato, K. R., Andras, J., Angenent, L. T., Antonopoulos, D. A., Apprill, A., Armitage, D., Ballantine, K., Bárta, J., Baum, J. K., Berry, A., Bhatnagar, A., Bhatnagar, M., Biddle, J. F., Bittner, L., Boldgiv, B., Bottos, E., Boyer, D. M., Braun, J., Brazelton, W., Brearley, F. Q., Campbell, A. H., Caporaso, J. G., Cardona, C., Carroll, J., Cary, S. C., Casper, B. B., Charles, T. C., Chu, H., Claar, D. C., Clark, R. G., Clayton, J. B., Clemente, J. C., Cochran, A., Coleman, M. L., Collins, G., Colwell, R. R., Contreras, M., Crary, B. B., Creer, S., Cristol, D. A., Crump, B. C., Cui, D., Daly, S. E., Davalos, L., Dawson, R. D., Defazio, J., Delsuc, F., Dionisi, H. M., Dominguez-Bello, M. G., Dowell, R., Dubinsky, E. A., Dunn, P. O., Ercolini, D., Espinoza, R. E., Ezenwa, V., et al. (2017) 'A communal catalogue reveals Earth's multiscale microbial diversity', *Nature,* 551(7681), 457-463.

Vorholt, J. A. (2012) 'Microbial life in the phyllosphere', *Nature Reviews Microbiology,* 10(12), 828-840.

Vásquez, A. and Olofsson, T. C. (2009) 'The lactic acid bacteria involved in the production of bee pollen and bee bread', *Journal of Apicultural Research,* 48(3), 189-195.

Wang, N., Qian, W., Suppanz, I., Wei, L., Mao, B., Long, Y., Meng, J., Müller, A. E. and Jung, C. (2011) 'Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the FRIGIDA homologue BnaA.FRI.a', *Journal of Experimental Botany,* 62(15), 5641-5658.

Wang, Q., Garrity, G. M., Tiedje, J. M. and Cole, J. R. (2007) 'Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy', *Applied and environmental microbiology,* 73(16), 5261-5267.

Wang, Y., Jorda, M., Jones, P. L., Maleszka, R., Ling, X., Robertson, H. M., Mizzen, C. A., Peinado, M. A. and Robinson, G. E. (2006) 'Functional CpG methylation system in a social insect', *Science,* 314(5799), 645-647.

White TJ, Bruns T, Lee S, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, New York, pp 315–322

Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

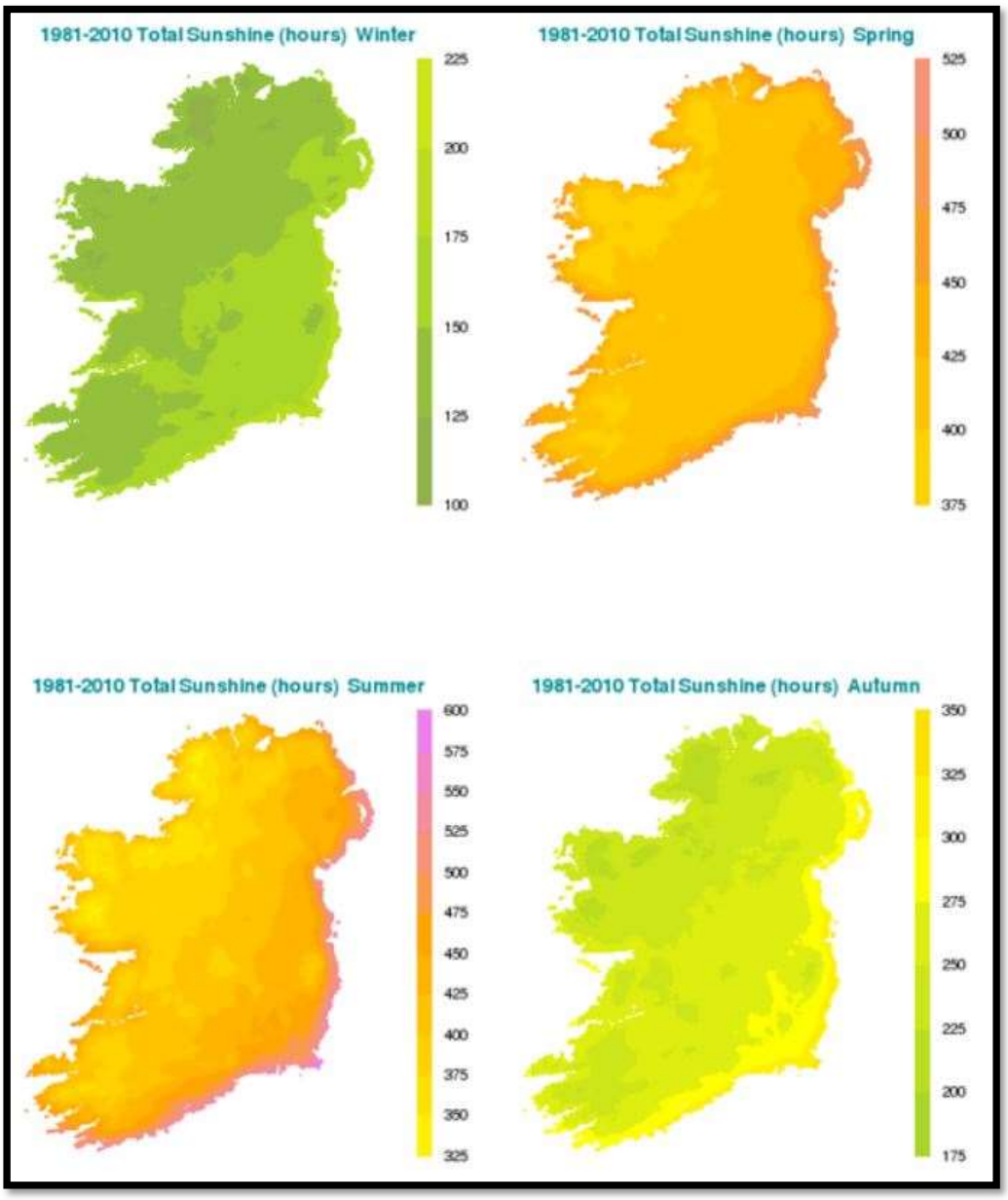
Wolf, A. A., Zavaleta, E. S. and Selmants, P. C. (2017) 'Flowering phenology shifts in response to biodiversity loss', *Proceedings of the National Academy of Sciences,* 114(13), 3463.

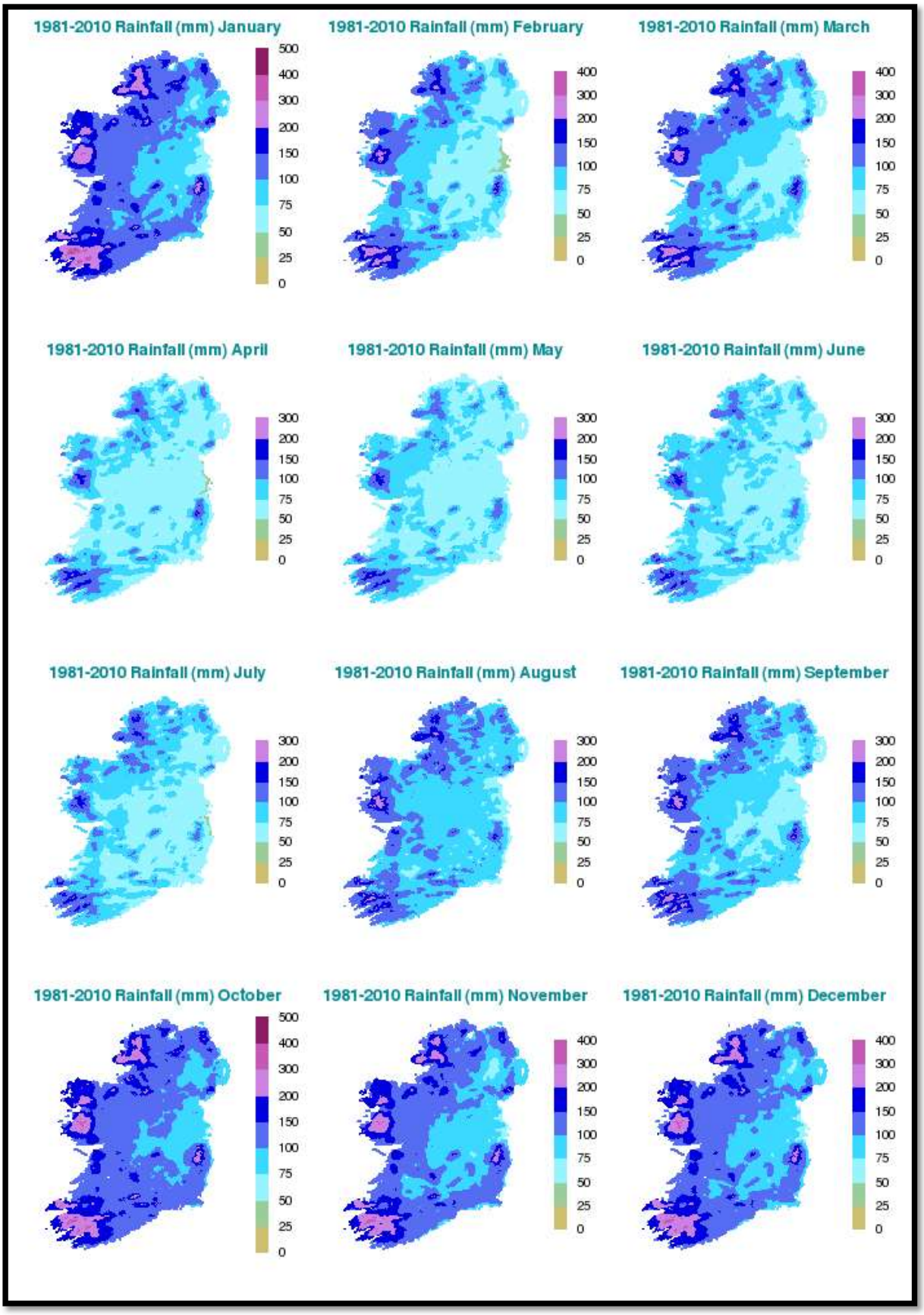
Zheng, H., Perreau, J., Powell, J. E., Han, B. F., Zhang, Z. J., Kwong, W. K., Tringe, S. G. and Moran, N. A. (2019) 'Division of labor in honey bee gut microbiota for plant polysaccharide digestion', *Proceedings of the National Academy of Sciences of the United States of America,* 116(51), 25909-25916.

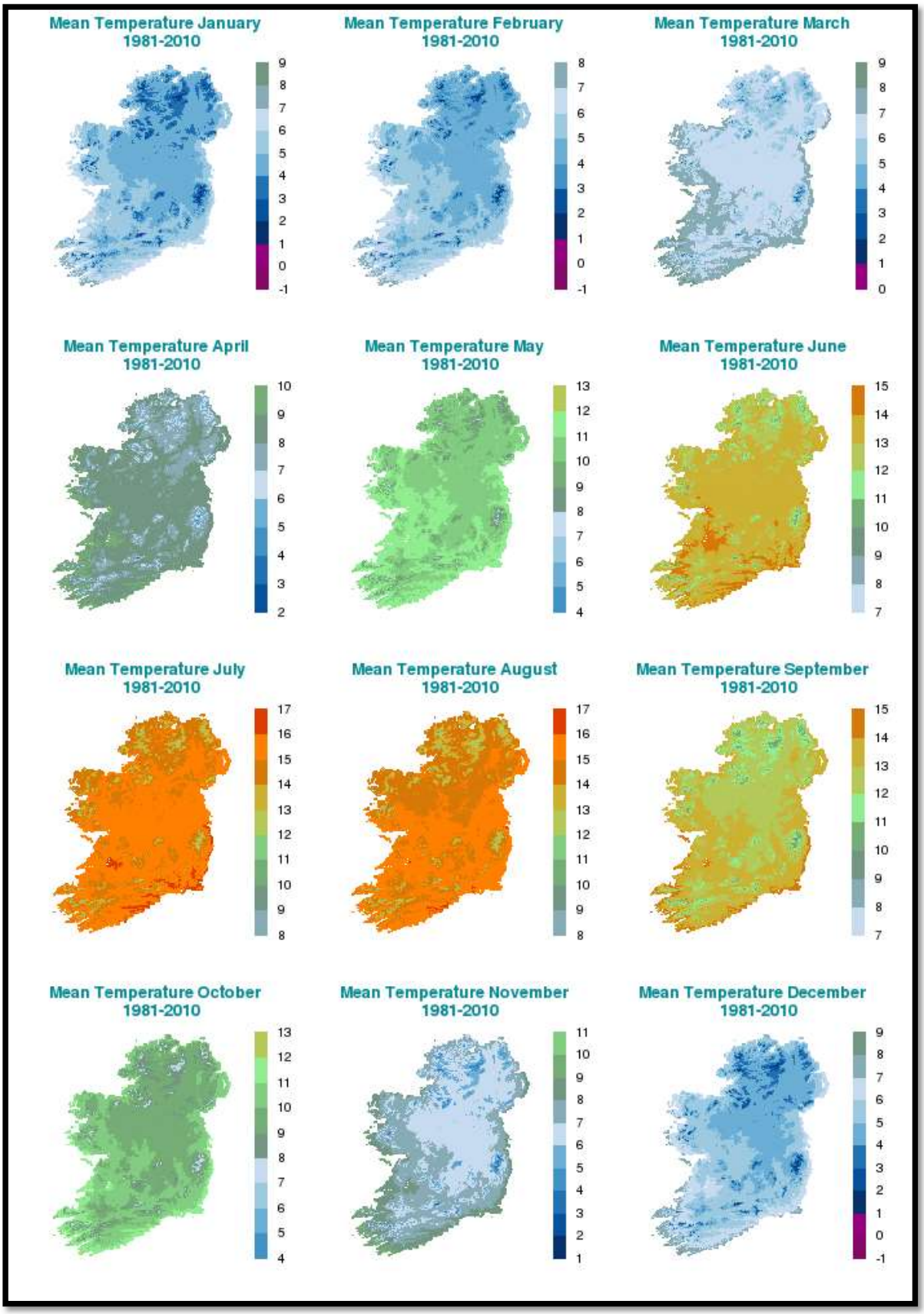
Zhu, K., Liu, M., Fu, Z., Zhou, Z., Kong, Y., Liang, H., Lin, Z., Luo, J., Zheng, H., Wan, P., Zhang, J., Zen, K., Chen, J., Hu, F., Zhang, C.-Y., Ren, J. and Chen, X. (2017) 'Plant microRNAs in larval food regulate honeybee caste development', *PLOS Genetics,* 13(8), e1006946.

Chapter 3 appendices I to IX

APPENDIX I – Climatic variables: Maps showing the mean annual total sunshine hours, monthly rainfall and monthly temperatures for the island of Ireland calculated from Met Eireann data from 1981 to 2010. All figures used with permission of Met Eireann.




1981-2010 Rainfall (mm) January
500
400
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) February
400
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) March
400
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) April
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) May
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) June
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) July
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) August
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) September
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) October
500
400
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) November
400
300
200
150
100
75
50
25
0
1981-2010 Rainfall (mm) December
400
300
200
150
100
75
50
25
0


Mean Temperature January
1981-2010
9 8 7 6 5 4 3 2 1 0 -1
Mean Temperature February
1981-2010
8 7 6 5 4 3 2 1 0 -1
Mean Temperature March
1981-2010
9 8 7 6 5 4 3 2 1 0
Mean Temperature April
1981-2010
10 9 8 7 6 5 4 3 2
Mean Temperature May
1981-2010
13 12 11 10 9 8 7 6 5 4
Mean Temperature June
1981-2010
15 14 13 12 11 10 9 8 7
Mean Temperature July
1981-2010
17 16 15 14 13 12 11 10 9 8
Mean Temperature August
1981-2010
17 16 15 14 13 12 11 10 9 8
Mean Temperature September
1981-2010
15 14 13 12 11 10 9 8 7
Mean Temperature October
1981-2010
13 12 11 10 9 8 7 6 5 4
Mean Temperature November
1981-2010
11 10 9 8 7 6 5 4 3 2 1
Mean Temperature December
1981-2010
9 8 7 6 5 4 3 2 1 0 -1

APPENDIX II - Initial hive assay

An initial pollen sampling and health assessment was trialled but proved to be overambitious and is included here as an example of such.

Pollen was taken by random sampling from frames containing stored beebread on brood box comb. The same sample point involved the removal of ≥25 young bees and ~10 stage five larvae along with an assessment of the colony pathogen load (*Varroa destructor*, Overt signs of deformed wing virus and tracheal mite, European foulbrood (EFB), American foulbrood (AFB) and Chalkbrood).

SAMPLING METHODS

1. Pollen

15mm lengths of standard diameter (3mm) drinks straws in a pre-determined set formation (Figure P1) were used to remove beebread samples from storage cells in an outer brood box frame. The straw was inserted gently to the base of the pollen cell, twisted and removed without damaging the back of the cell wall

Sample size: 12 straws per frame
4 frames per colony from 2 regions (2 from centre of brood area and 2 from peripheral frames)

2. Health assessment

ON SITE

a. varroa: Sugar shaker assessment per the NIHBS/NUIG guidelines

b. AFB/EFB : 10 minute visual assessment of random brood frames. Suspect cells will be removed for confirmation.

c. Overt signs : 10 minute visual count on random brood frames with removal of any K-wing and deformed wing bees to prevent them being counted twice and for microscopic analysis.

d. Chalkbrood : 5 minute visual count of open cells from randomly chosen brood frames

OFF SITE

e. Nosema ; 25 bees from each hive, each visit. Light microscopy identification from bee gut. Nosema apis will be distinguished from N. ceranae using molecular methods.

f. Tracheal mite: Light microscopy of the tracheae of k-wing bees that are removed.

g. Viruses : Viral presence/absence– Laboratory based molecular discrimination method.

h. Gut flora: Gut section from 10 of the 25 bees which were used for Nosema assessment.

i. Cuticules of larvae: 10 pre-pupa stage larvae from each hive, each visit. Their cuticule will be assessed chemically for differences between healthy and unhealthy colonies based on the other assessments listed above.

Figure P1 – Pollen sampling frame consisting of a net matrix equivalent to the average number of cells on a brood frame with 12 pre-marked sample point holes. The frame overlays the beebread storage cells and a straw is used to remove a sample of the beebread in the indicated cell (or nearest full cell). The pre-marked holes remain the same between hives.

The beebread sampling method used a pre-marked randomised selection template which proved to be very time consuming in its execution. varroa load was assessed used the sugar shake method and the on-site pathogen assessment was a visual check for the disease signs on the brood frames. Conducted together, these methods required the hives to be open for a period which overstressed both the colonies and consequently the person sampling. It is likely that the stress on the sampler would have coerced them to take shortcuts, affecting randomisation protocols and accuracy. Furthermore, as it may have affected colony behaviour, any future sampling would be compromised. The final methods as used placed a high priority on minimising stress to both the hive and the sample taker.

i)

ii)

Figure P2 – i) Sampling straws in place with the net matrix removed ii) The straws were rotated between the fingers before removal with the beebread sample inside.

APPENDIX III - Whole DNA extraction pipeline

A weighing scale accurate to four decimal places is necessary.

Weighing and homogenisation of the unsorted (pollen plus detritus) sample

1. Weigh the unsorted sample including the bag having first removed any external ice

2. Allow the sample to defrost for approximately one hour.

3. Empty the contents of the bag into a clean container. Add a recorded volume of sufficient distilled water to cover the sample and leave to defrost. Record the water volume used and leave overnight to defrost and saturate the sample.

4. Weigh the empty bag and calculate the net unsorted sample weight.

5. If required, add a further recorded volume of distilled water to help suspend the pollen grains. (The volume of water used should be kept to a minimum.)

6. Mix on a stirrer plate at moderate speed for approximately 10 minutes or until all the pollen pellets are broken down and the mix looks homogenised.

Detritus filtering

7. To remove detritus, cut and weigh a piece of 500µm mesh. Filter the mixture through the mesh into a clean container. Record the volume of distilled water used to wash the pollen through the mesh.

8. Place a clean stirrer magnet in the mix and stir until the suspension is circulating gently.

A- Samples for future use

9. Using a pipette take six separate samples from the middle of the circulating suspension and place them into separate 1.5mL microtubes.

i) Pollen density calculation - Seal 2 samples with parafilm to prevent moisture loss.

ii) Future microscopy ID – Centrifuge 2 samples at maximum rpm for 3 minutes. Pipette off and discard the supernatant.

iii) Future DNA extraction - Centrifuge 2 samples at maximum for 3 minutes. Pipette off and discard the supernatant. Add 1000uL of 100% Ethanol, vortex until all the pollen is in suspension.

10. Store all six samples at -20C in light tight containers

B- Samples for immediate downstream processing

Producing a 50mg sub-sample marker

(NB: Beforehand, weigh ≥10 of the screwtop microtubes intended for use and calculate the mean microtube weight)

1 Pipette 1000uL of the circulating suspension into a 2mL screwtop microtube

2 Centrifuge at maximum for 3 minutes

3 Carefully pipette off the supernatant.

4 Weigh the microtube and sub-sample and deduct the mean tube weight to determine the pollen weight.

Due to the different suspended pollen concentrations the addition or removal of sample material may be necessary.
If the sample weight is <50mg, add additional volumes of the suspension to the microtube and repeat steps 2 to 4 until 50mg of sample is obtained.
If the sample weight is >50mg, remove material as necessary and centrifuge for 30 secs.

DNA extraction using Machery-Nagel Nucleospin food kit and Spexprep Mini-G bead beater

1. To the 50mg pollen sample add:
    80mg of acid washed sand
    38 mg of 200um Zircomium beads
    85mg of 400um Zirconium Beads
    97mg of 800um Zr. Beads
    Two 2.8mm ceramic beads
2. Add 550uL CF buffer preheated to 65^{o}C
3. Disrupt in the bead beater at 1500rpm for 4 mins
4. Add 10uL Proteinase K solution
5. Incubate at 65^{o}C for 60mins.
6. Centrifuge at 1000rpm for 2 mins
7. Transfer the supernatant into a fresh 1.5mL microtube
8. To the supernatant, add 1 equal volume of C4 buffer
9. Add 1 equal volume 96-100% ethanol
10. Pipette a maximum of 750uL of the sample to the Nucleospin column in a collection tube.
11. Centrifuge at maximum for 1 min
    Repeat steps 10 and 11 until all the sample has been centrifuged in the column.
    Discard the flow-through
12. Pipette 400uL CQW buffer onto the column and centrifuge at maximum for 1 min
    Discard the flow-through
13. Pipette 680uL of C5 buffer onto the column and centrifuge at maximum for 1 min
    Discard the flow-through
14. Pipette another 220uL C5 buffer onto the column and centrifuge at maximum for 1 min
    Discard the flow-through
15. Place the column in a clean 1.5mL microtube. Pipette 100uL of elution buffer CE (preheated to 70^{o}C) onto the column. Incubate at room temp for 5 min. Centrifuge at maximum for 1 min.
16. Discard the column. Store the eluted DNA at -20 ^{o}C until required.

APPENDIX IV - Sample reads, Post QC reads and distinct sequences (arranged by the sample name to compliment Table 3 in the main text) achieved after read quality control using a Phred Q value of 25 (99.7%) and a maximum expected error of 1. The sample codes eg: H4.8.6 contain the Hive name (H4), Month (August) and Year (201**6**). These data are also arranged by sample number in appendix IV. 'E' before a month such as E9 means the sample was taken early in that month, in this case, early September.

Sample	R1 & R2 Reads	Post-quality control R1 reads	Distinct sequences	Number of pollen types	>45%	>15% to 45%	>3% to 15%	1% to 3%	<1%
C2.10.6	69209	8212	2602	17	1			1	15
C2.12.6	49087	18927	12477	25		4	1	2	18
C2.3.7	21502	13725	12740	13	1	2	1		9
C2.4.7	19139	11486	9632	19	1	1	2		15
C2.5.7	8163	4061	3311	14	1	2		2	9
C2.6.7	5982	2065	1786	18	1	1	2	1	13
C2.7.6	26284	13468	9859	22		4	2		16
C2.8.6	19976	229	131	12		2	4	4	2
C2.9.6	16311	4705	4643	8	1				7
C2.E9.6	4946	230	82	9	1	1	2	5	
C3.10.6	43394	24357	19893	13	1				12
C3.12.6	33553	13412	12775	25		3	4		18
C3.3.7	19832	13127	12124	14	1	2		1	10
C3.4.7	20796	10346	8850	15	1	2	1	1	10
C3.5.7	2906	340	212	8	1	2		3	2
C3.6.7	19439	10614	9936	13	1	1		2	9
C3.7.7	26386	17208	16512	19	2				17
C3.8.6	32007	5634	3608	23		3	1	3	16
C3.9.6	44658	16719	10063	12	1			1	10
C3.E9.6	6047	144	70	7	1	1	1	4	
C8.10.6	47506	27528	27211	16	1				15
C8.12.6	62304	13701	11410	21		2	2	3	14
C8.3.7	12493	8655	7986	9	1	2			6
C8.4.7	14876	7947	5827	14	1	1	1	1	10
C8.5.7	42834	7203	5978	17		3	3	1	10
C8.6.7	3380	1005	877	7	1		2		4
C8.7.6	33423	12439	10568	21	1	1	2		17
C8.7.7	16673	3657	2068	16	1	1		5	9
C8.8.6	11398	2065	944	22		3	3	1	15
C8.9.6	15989	8222	7042	7	1				6
C8.E9.6	5945	108	40	10		3	3	4	
H14.4.7	47228	26969	23571	18	1	1	1	1	14
H14.5.7	10397	3552	2980	17	1	2	2	4	8
H14.7.7	18352	10708	7826	23		3		3	17
H14.8.6	37443	9482	6945	23	1		1		21
H4.4.7	14880	7390	6097	20	1	1	2	1	15
H4.5.7	10569	899	419	17		3	1	6	7
H4.7.7	6415	3846	3541	12	1	2		1	8
H4.8.6	50455	28991	14312	26	1	1		1	23
H4.9.6	35830	22247	20941	14	1				13
H5.4.7	7188	229	123	10	1	1	2	3	3
H5.5.7	3668	54	45	8	1		3	4	
H5.7.7	20396	5675	5136	19	1	1	2		15
H5.8.6	1817	156	49	9		2	5	2	
H5.9.6	44139	25910	23713	16	1	1	1		13
JH17.4.7	27767	17628	15272	15	1	1	1	1	11
JH17.6.7	564	64	23	5	1	1	3		
JH17.7.7	13244	4253	3861	14		3	2		9
JH17.8.6	27025	16695	8045	18	1	1	1	2	13
JH17.9.6	22767	13325	11692	11	1		2		8
JH6.4.7	14278	8421	8040	18	1	1	1	1	14
JH6.7.7	12093	6079	5159	16		3	1		12

JH6.8.6	21785	11148	6489	18	1	2	1	1	13
JH6.9.6	44492	23899	18138	15	1	1	1		13
LG2.10.6	28836	11709	11507	8	1				7
LG2.4.7	20588	13817	12954	13	1	1	1	1	9
LG2.6.7	5214	808	522	19	1	1	4	2	11
LG2.7.7	16361	10069	7916	20	1	1	1	1	16
LG2.8.6	295	119	99	8		3	2	3	
LG2.9.6	7046	2877	2792	8	1				7
LG2.E4.7	12764	7392	6298	11	1	1		2	7
M1.5.7	22763	11434	8742	19	1		4	2	12
M1.7.7	27128	14531	12409	22	1		3	1	17
M1.8.6	43025	25921	17061	27	1	1	2	1	22
M1.9.6	44389	27300	23962	13	2				11
M2.5.7	21712	9872	7390	25		1	6	3	15
M2.7.7	18577	11447	9602	12	1	1	2	1	7
M2.8.6	46676	22511	17181	27	1	1			25
M2.9.6	46351	28597	26295	19	1	1			17
M3.5.7	20033	9768	7223	27	1	1	4	2	19
M3.7.7	36217	22095	19119	17	1	1		1	14
M3.8.6	1685	206	103	9	1	1	4	1	2
M3.9.6	21715	11536	11277	11	1	1			9
PD3.4.7	7052	4206	3971	17	1	1	3	1	11
PD3.7.7	22555	10453	10071	19	1	1		1	16
PD3.8.6	26172	4795	2626	10	1		1	1	7
PD3.9.6	29885	16278	15814	14	1		1	1	11
PD6.4.7	17917	10995	8041	9		3	2	1	3
PD6.7.7	22926	5949	5632	19	1	1	1	3	13
PD6.8.6	25627	2939	247	12	1	1	1	4	5
PD6.9.6	18013	10395	10267	8	1				7
PD9.4.7	13216	7593	7162	18		2	4		12
PD9.7.7	16985	353	95	15		1	9	5	
PD9.8.6	21249	5391	1379	10	1		1	2	6
PD9.9.6	32696	10401	6929	13	1			2	10
SK212.4.7	14634	9919	9092	23		3	1		19
SK212.7.7	16426	275	197	13	1		8	3	1
SK212.8.6	44714	21228	17321	18	1	1	4		12
SK212.9.6	31231	20609	8923	12		2	2	1	7
SK42.4.7	12993	7682	7049	18	1	1	3		13
SK42.7.7	21800	13058	11989	14	1	1	1	2	9
SK42.8.6	31823	17590	13735	21		2	5	1	13
SK42.9.6	38728	23361	19398	12		3	1		8
SK85.4.7	32521	14326	10118	20	1	2		1	16
SK85.7.7	31907	974	7641	20	1	2	2	1	14
SK85.8.6	42819	20737	16451	18		3	3	1	11
Totals	**2264494**	**1010675**	**820274**	**1496**	**74**	**124**	**156**	**128**	**1015**

APPENDIX V – Identified pollen species with one common name used in Ireland along with the status and flowering period. Synonyms for both scientific and common names should be checked by the user of this information.

Full scientific name	Common name	Status	Flowering period
Acer platanoides	Norway maple	Introduced	J F **M A M** J J A S O N D
Acer pseudoplatanus	Sycamore	Introduced	J F M A **M J** J A S O N D
Aesculus carnea	Red horse chestnut	Cultivar	J F M A **M** J J A S O N D
Allium ampeloprasum	Wild leek	Introduced	J F M A M **J J A S** O N D
Alnus glutinosa	Common alder	Native	J **F M A** M J J A S O N D
Berberis aquifolium	Oregan grape	Cultivar	J **F M A** M J J A S O N D
Betula pendula	Silver birch	Native	J F M **A M** J J A S O N D
Borago officinalis	Borage	Cultivar	J F M **A M J J A S** O N D
Brassica napus	Rapeseed	Cultivar	J F M **A M J J A** S O N D
Brassica oleracea	Wild cabbage	Introduced	J F M A **M J J A** S O N D
Calluna vulgaris	Ling heather	Native	J F M A M J **J A S O** N D
Centaurea cyanus	Cornflower	Introduced	J F M A **M J J A S O** N D
Centaurea nigra	Common knapweed	Native	J F M A M **J J A S O** N D
Chamerion angustifolium	Rosebay willowherb	Invasive	J F M A M **J J A S** O N D
Cichorium intybus	Chicory	Introduced	J F M A M **J J A S** O N D
Cirsium arvense	Field thistle	Native	J F M A M **J J A S** O N D
Clematis vitalba	Travellers joy	Cultivar	J F M A M J **J A S** O N D
Convolvulus arvensis	Field bindweed	Native	J F M A M **J J A S** O N D
Crataegus monogyna	Hawthorn	Native	J F M A **M J** J A S O N D
Cytisus scoparius	Broom	Native	J F M **A M J** J A S O N D
Echium vulgare	Viper's bugloss	Native	J F M A **M J J A S** O N D
Epilobium angustifolium	Rosebay willowherb	Invasive	J F M A M **J J A S** O N D
Erica cinerea	Bell Heather	Native	J F M A M **J J A S** O N D
Erysimum alaicum	None	Cultivar	J F M **A M J** J A S O N D
Erysimum canum	None	Cultivar	J F M **A M J** J A S O N D
Erysimum cheiri	Wallflower	Cultivar	J F M **A M J** J A S O N D
Erysimum virescens	None	Cultivar	J F M **A M J** J A S O N D
Fagus sylvatica	European beech	Native	J F M **A M** J J A S O N D
Ficaria verna	Lessor celandine	Native	**J F M A M** J J A S O N **D**
Filipendula ulmaria	Meadowsweet	Native	J F M A M **J J A S** O N D
Hedera helix	Ivy	Native	J F M A M J J A **S O N** D
Heracleum sphondylium	Hogweed	Native	J F M A M **J J A S** O N D
Ilex aquifolium	Holly	Native	J F M A **M J J A** S O N D
Ligustrum vulgare	Privet	Native	J F M A **M J J** A S O N D
Lotus corniculatus	Bird's-foot trefoil	Native	J F M A M **J J A S** O N D
Lythrum salicaria	Purple-loosestrife	Native	J F M A M **J J A S** O N D
Malus pumila	Apple, cultivated	Cultivar	J F M **A M** J J A S O N D
Melilotus albus	White melilot	Cultivar	J F M A M **J J A** S O N D
Onobrychis viciifolia	Sainfoin	Introduced	J F M A M **J J A** S O N D
Origanum vulgare	Wild marjoram	Native	J F M A M J **J A S** O N D
Papaver rhoeas	Common poppy	Introduced	J F M A M **J J A** S O N D
Plantago lanceolata	Ribwort plantain	Native	J F M **A M J J A S O** N D

Populus euphratica	Euphrates poplar	Introduced	J **F M A** M J J A S O N D
Populus trichocarpa	California poplar	Introduced	J **F M A** M J J A S O N D
Prunus avium	Wild cherry	Native	J F M **A M** J J A S O N D
Prunus cerasifera	Cherry plum	Cultivar	J **F M** A M J J A S O N D
Prunus dulcis	Almond	Cultivar	J F **M A** M J J A S O N D
Prunus laurocerasus	Cherry laurel	Invasive	J F M **A M J** J A S O N D
Prunus spinosa	Blackthorn	Native	J F **M A M** J J A S O N D
Quercus robur	Pedunculate oak	Native	J F M **A M** J J A S O N D
Ranunculus bulbosus	Bulbous buttercup	Native	J F **M A M J J** A S O N D
Raphanus raphanistrum	Wild radish	Introduced	J F M **A M J J A S** O N D
Robinia pseudoacacia	Black locust	Cultivar	J F M A **M J** J A S O N D
Rosa canina	Dog rose	Native	J F M A M **J J A** S O N D
Rubus fruticosus	Blackberry	Native	J F M A **M J J A S** O N D
Rubus idaeus	Raspberry	Native	J F M A M **J J** A S O N D
Rubus plicatus	Blackberry	Native	J F M A **M J J A S** O N D
Salix caprea	Goat willow	Native	J F **M A** M J J A S O N D
Sambucus nigra	Elder	Native	J F M A **M J** J A S O N D
Sambucus palmensis	Elder spp	Introduced	J F M A **M J** J A S O N D
Sinapis arvensis	Charlock	Introduced	J F M **A M J J A S O** N D
Syringa vulgaris	Lilac	Cultivar	J F M A **M** J J A S O N D
Taraxacum officinale	Dandelions	Native	J F **M A M J J A S O** N D
Tilia americana	American lime	Introduced	J F M A M **J** J A S O N D
Trifolium pratense	Red clover	Native	J F M A **M J J A S O** N D
Trifolium repens	White clover	Native	J F M A M **J J A S** O N D
Ulex europaeus	Gorse	Native	**J F M A M J J A S O N D**
Vicia faba	Broad bean	cultivar	J F M A **M J** J A S O N D

APPENDIX VI – Identified species arranged by flowering season (in red). The ranks refer to the relative abundance of each species from all samples (N=96) Species status' (in Ireland) and flowering seasons were obtained from www.wildflowersofireland.net, www.irishwildflowers.ie and www.treeandlandscape.ie.

Taxonomic name	Common name	Status	Rank	Flowering
Ficaria verna	Lessor celandine	Native	64	**J F M A M** J J A S O N **D**
Ulex europaeus	Gorse	Native	5	**J F M A M J J A S O N D**
Prunus cerasifera	Cherry plum	Cultivar	65	J **F M** A M J J A S O N D
Berberis aquifolium	Oregon grape	Cultivar	21	J **F M A** M J J A S O N D
Populus euphratica	Euphrates poplar	Introduced	55	J **F M A** M J J A S O N D
Populus trichocarpa	California poplar	Introduced	56	J **F M A** M J J A S O N D
Alnus glutinosa	Common alder	Native	68	J **F M A** M J J A S O N D
Salix caprea	Goat willow	Native	31	J F **M A** M J J A S O N D
Prunus dulcis	Almond	Cultivar	46	J F **M A** M J J A S O N D
Acer platanoides	Norway maple	Introduced	51	J F **M A M** J J A S O N D
Prunus spinosa	Blackthorn	Native	59	J F **M A M** J J A S O N D
Ranunculus bulbosus	Bulbous buttercup	Native	10	J F **M A M J J** A S O N D
Taraxacum officinale	Dandelions	Native	19	J F **M A M J J A S O** N D
Prunus avium	Wild cherry	Native	37	J F M **A M** J J A S O N D
Fagus sylvatica	European beech	Native	39	J F M **A M** J J A S O N D
Quercus robur	Pedunculate oak	Native	45	J F M **A M** J J A S O N D
Betula pendula	Silver birch	Native	53	J F M **A M** J J A S O N D
Malus pumila	Apple, cultivated	Cultivar	54	J F M **A M** J J A S O N D
Erysimum virescens	None	Cultivar	43	J F M **A M J** J A S O N D
Erysimum alaicum	None	Cultivar	44	J F M **A M J** J A S O N D
Cytisus scoparius	Broom	Native	41	J F M **A M J** J A S O N D
Prunus laurocerasus	Cherry laurel	Invasive	52	J F M **A M J** J A S O N D
Erysimum canum	None	Cultivar	57	J F M **A M J** J A S O N D
Erysimum cheiri	Wallflower	Cultivar	67	J F M **A M J** J A S O N D
Brassica napus	Rapeseed	Cultivar	1	J F M **A M J J A** S O N D
Raphanus raphanistrum	Wild radish	Introduced	26	J F M **A M J J A S** O N D
Borago officinalis	Borage	Cultivar	49	J F M **A M J J A S** O N D
Sinapis arvensis	Charlock	Introduced	4	J F M **A M J J A S O** N D
Plantago lanceolata	Ribwort plantain	Native	22	J F M **A M J J A S O** N D
Aesculus carnea	Red horse chestnut	Cultivar	63	J F M A **M** J J A S O N D
Syringa vulgaris	Lilac	Cultivar	66	J F M A **M** J J A S O N D
Acer pseudoplatanus	Sycamore	Introduced	8	J F M A **M J** J A S O N D
Sambucus nigra	Elder	Native	9	J F M A **M J** J A S O N D
Vicia faba	Broad bean	cultivar	12	J F M A **M J** J A S O N D
Crataegus monogyna	Hawthorn	Native	20	J F M A **M J** J A S O N D
Sambucus palmensis	Elder spp	Introduced	60	J F M A **M J** J A S O N D
Robinia pseudoacacia	Black locust	Cultivar	62	J F M A **M J** J A S O N D
Ligustrum vulgare	Privet	Native	58	J F M A **M J J** A S O N D
Brassica oleracea	Wild cabbage	Introduced	15	J F M A **M J J A** S O N D
Ilex aquifolium	Holly	Native	27	J F M A **M J J A** S O N D
Rubus plicatus	Blackberry	Native	17	J F M A **M J J A S** O N D

Rubus fruticosus	Blackberry	Native	23	J F M A **M J J A S** O N D

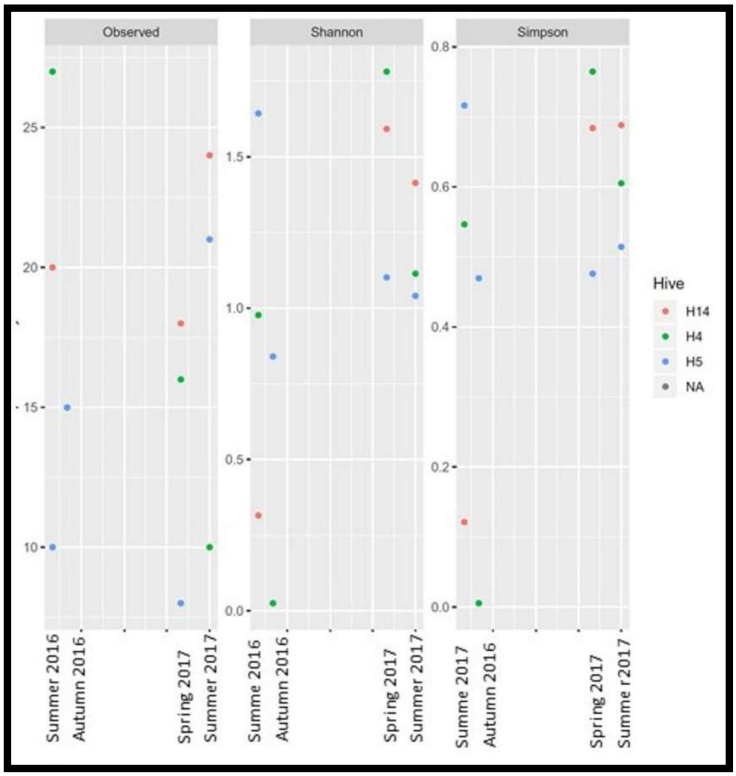
Continued…

Taxonomic name	Common name	Status	Rank	Flowering
Echium vulgare	Viper's bugloss	Native	50	J F M A **M J J A S** O N D
Centaurea cyanus	Cornflower	Introduced	35	J F M A **M J J A S O** N D
Trifolium pratense	Red clover	Native	36	J F M A **M J J A S O** N D
Tilia americana	American lime	Introduced	61	J F M A M **J** J A S O N D
Rubus idaeus	Raspberry	Native	3	J F M A M **J J** A S O N D
Rosa canina	Dog rose	Native	13	J F M A M **J J A** S O N D
Papaver rhoeas	Common poppy	Introduced	18	J F M A M **J J A** S O N D
Onobrychis viciifolia	Sainfoin	Introduced	24	J F M A M **J J A** S O N D
Melilotus albus	White melilot	Cultivar	34	J F M A M **J J A** S O N D
Trifolium repens	White clover	Native	6	J F M A M **J J A S** O N D
Filipendula ulmaria	Meadowsweet	Native	7	J F M A M **J J A S** O N D
Lythrum salicaria	Purple-loosestrife	Native	11	J F M A M **J J A S** O N D
Cirsium arvense	Field thistle	Native	14	J F M A M **J J A S** O N D
Lotus corniculatus	Bird's-foot trefoil	Native	16	J F M A M **J J A S** O N D
Heracleum sphondylium	Hogweed	Native	25	J F M A M **J J A S** O N D
Cichorium intybus	Chicory	Introduced	28	J F M A M **J J A S** O N D
Chamerion angustifolium	Rosebay willowherb	Invasive	32	J F M A M **J J A S** O N D
Epilobium angustifolium	Rosebay willowherb	Invasive	33	J F M A M **J J A S** O N D
Erica cinerea	Bell Heather	Native	38	J F M A M **J J A S** O N D
Allium ampeloprasum	Wild leek	Introduced	47	J F M A M **J J A S** O N D
Convolvulus arvensis	Field bindweed	Native	48	J F M A M **J J A S** O N D
Centaurea nigra	Common knapweed	Native	30	J F M A M **J J A S O** N D
Origanum vulgare	Wild marjoram	Native	40	J F M A M J **J A S** O N D
Clematis vitalba	Travellers joy	Cultivar	42	J F M A M J **J A S** O N D
Calluna vulgaris	Ling heather	Native	29	J F M A M J **J A S O** N D
Hedera helix	Ivy	Native	2	J F M A M J J A **S O N** D

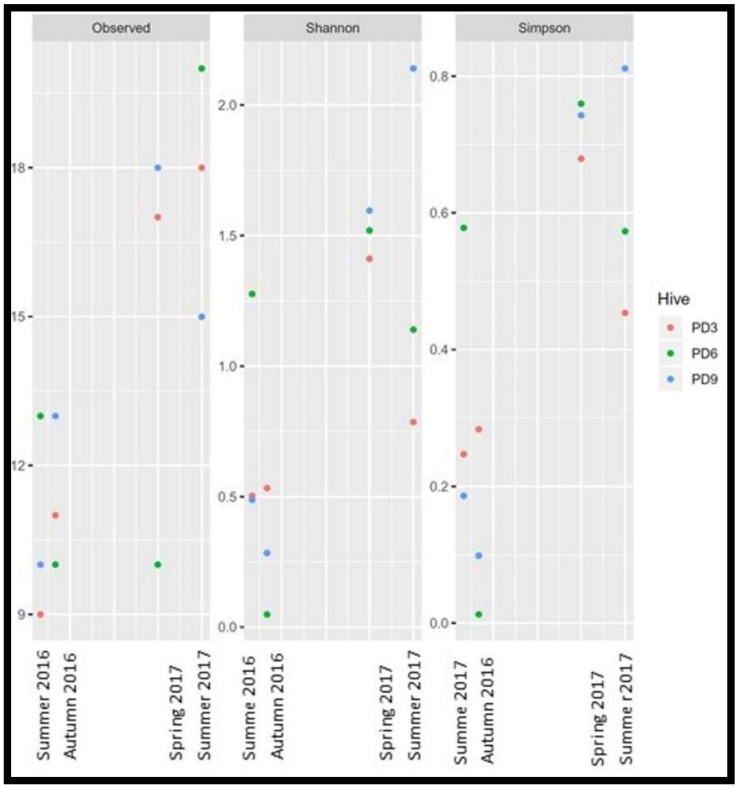
APPENDIX VII – Diversity plots

Plotted Observed, Simpson and Shannon diversity indices for each of the seven apiaries

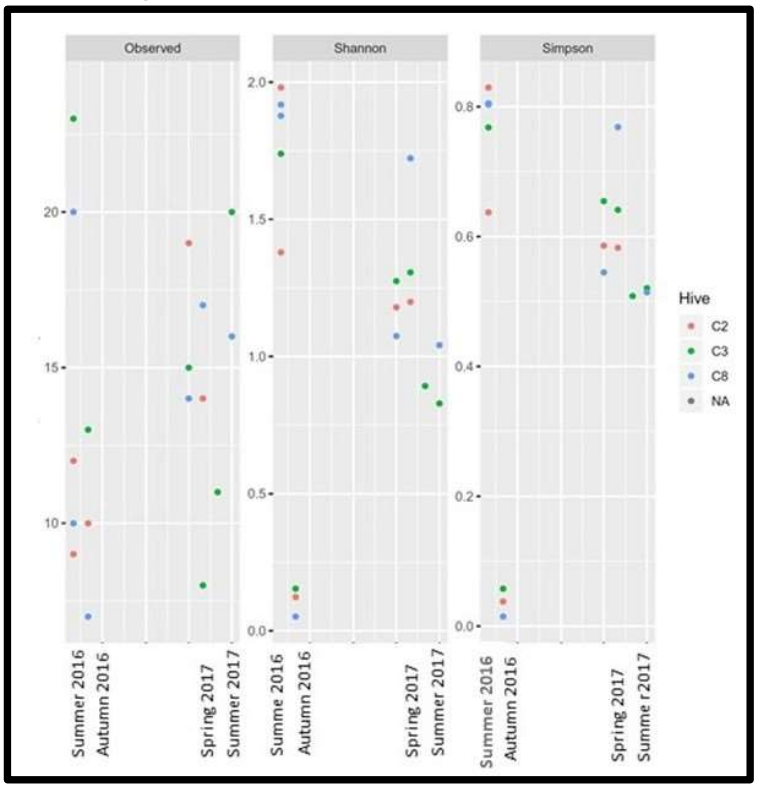
a) Coolmore hives H4, H5 and H14



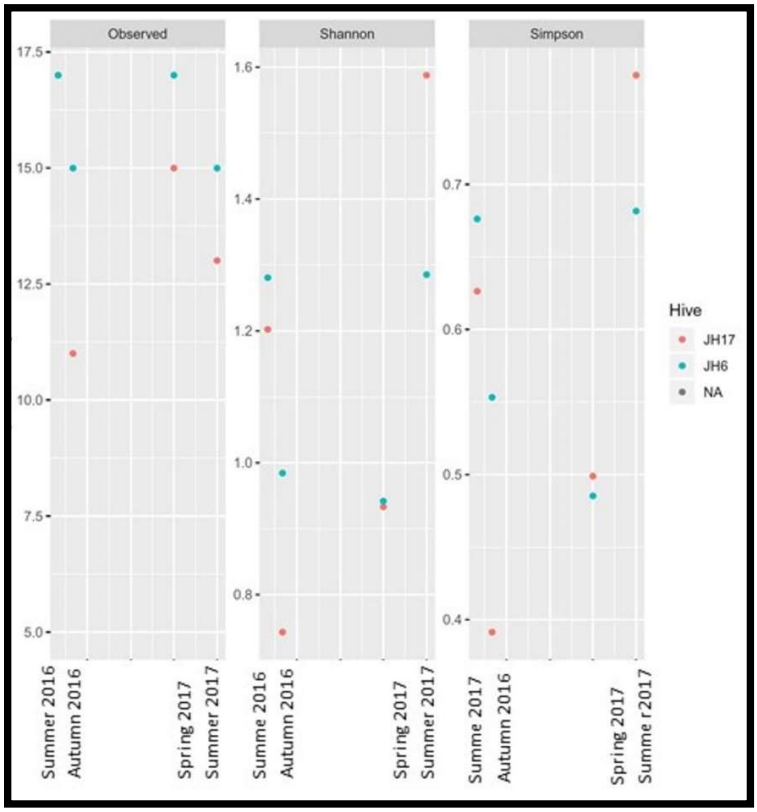
b) Youghal hives PD3, PD6 and PD9



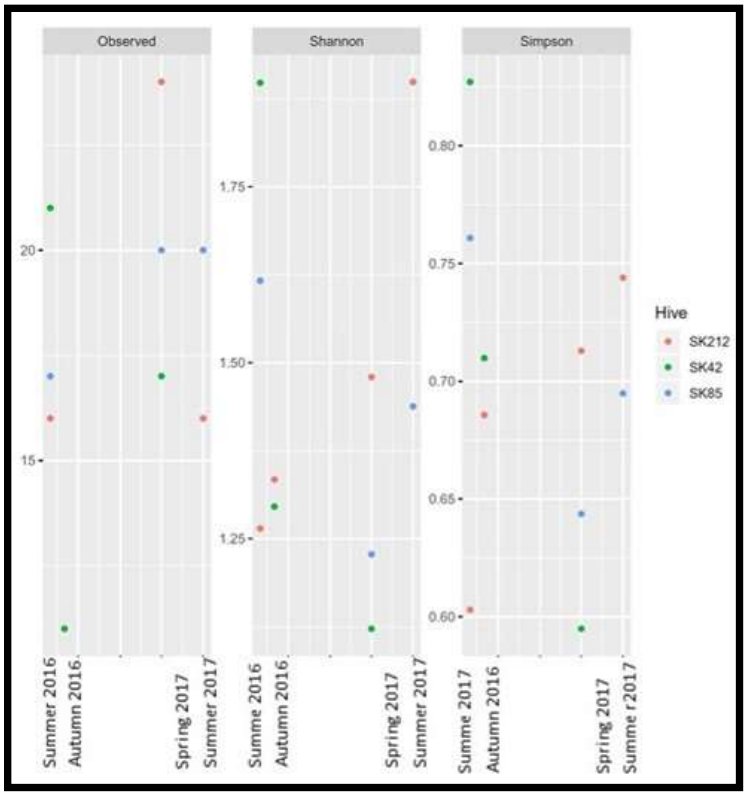
c) Dungarvan hives C2, C3 and C8



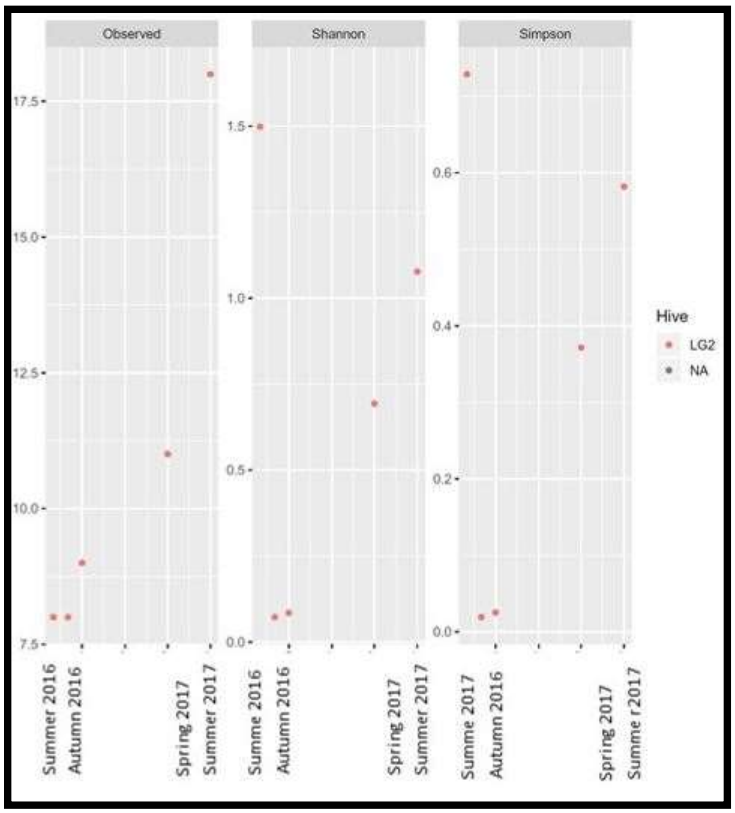
d) Campile hives JH6 and JH17



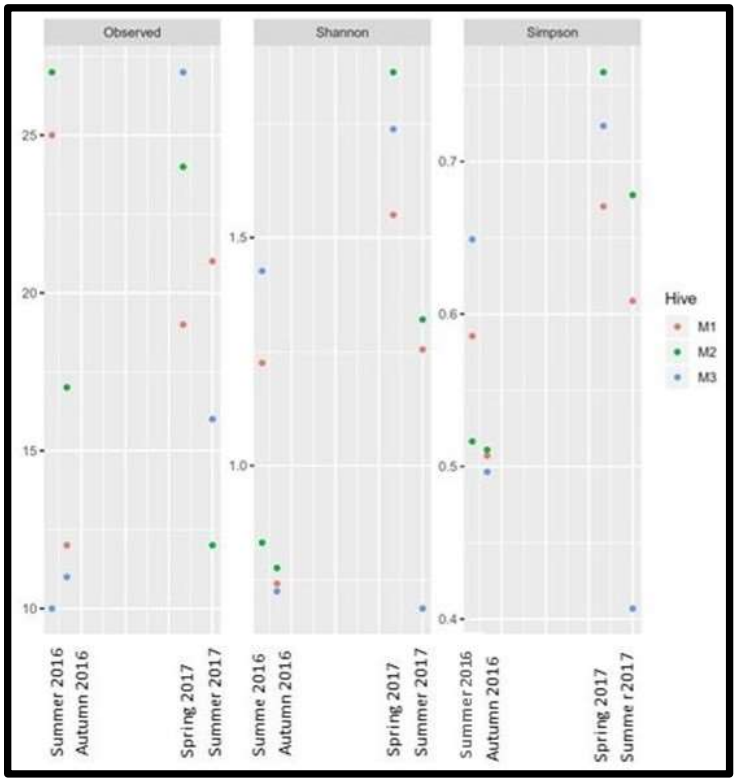
e) Fethard-on-sea hives SK42, SK85 and SK212



f) Hacketstown hive LG2



g) Labane hives M1, M2 and M3

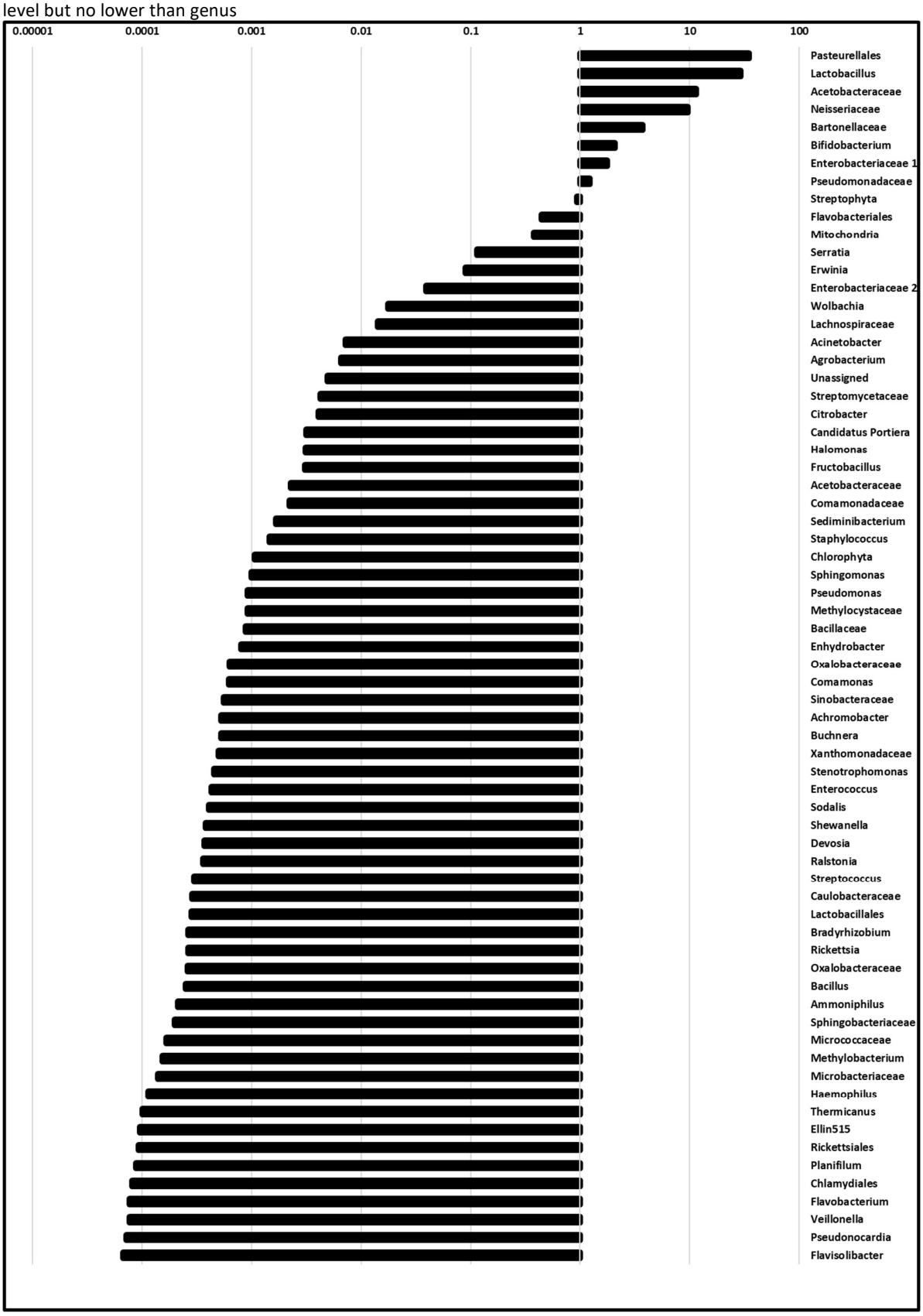


APPENDIX VIII– 16s throughput and OTUs per sample of the 4,314,380 paired-end reads following quality filtering steps QF1 and QF2 (see main text). Organized in descending order by the final throughput after QF2.

Sample ID	Reads after QF1	Reads after QF2	OTUs per sample
SK85.9.6	57286	47298	32
PD3.7.7	56614	46869	39
H4.9.6	53809	44082	45
JH17.9.6	50933	43637	33
SK42.9.6	41976	35441	52
LG2.9.6	32787	28700	36
JH17.4.7	35645	27935	47
JH17.5.7	31362	27858	27
PD6.4.7	31486	27473	30
JH17.6.6	32700	27047	32
C3.7.6	30879	26600	37
SK42.8.6	32212	26238	41
JH6.4.7	29570	26232	58
C2.5.7	31763	25970	55
PD3.4.7	33189	25884	37
C8.5.7	30418	24541	40
JH6.9.6	30831	24388	36
JH6.5.7	27350	24300	33
C3.9.6	29100	23813	42
SK212.8.6	28463	22992	54
C8.7.7	28053	22916	42
C2.E9.6	27081	22464	23
M2.5.7	27311	22162	34
C8.9.6	26452	21647	39
C3.5.7	26607	21552	40
JH6.6.6	24076	21388	36
H4.7.7	25500	20766	35
H5.7.7	24405	20226	38
LG2.7.7	24559	19966	48
H14.7.7	23962	19938	35
SK85.7.7	22501	19847	29
C3.7.7	23194	19095	28
SK85.8.6	21952	18674	31
C2.4.7	21953	18485	30
C2.8.6	21458	18346	33
M1.5.7	22180	18288	38
LG2.10.6	20309	17174	31
H4.8.6	20082	16926	36
PD9.4.7	19976	16916	30
JH6.10.6	19597	16640	29
LG2.5.7	19585	16485	38
H14.8.6	19561	16235	33
H5.9.6	19790	16159	33
M3.9.6	19340	16137	31
C2.7.7	19935	16083	32
PD6.7.7	19235	16072	33
C8.6.7	17998	15869	34
M3.5.7	19382	15772	31
PD6.8.6	18165	15418	30
C2.9.6	18847	15214	36
M2.7.7	18955	15053	45
SK85.4.7	17905	14976	32
C8.7.6	16775	14575	28

M2.9.6	18163	14528	29
C3.4.7	18010	14275	37
C2.7.6	16909	14162	32
PD9.7.7	16642	14074	32
SK42.4.7	15778	13237	38
JH17.10.6	15000	13197	30
M2.8.6	15751	13031	35
M1.8.6	15716	13003	36
LG2.4.7	16089	12473	33
PD6.9.6	15317	12418	33
SK212.9.6	15232	12300	32
H14.5.7	14727	12221	30
C8.4.7	15527	12203	26
PD3.9.6	14942	12169	29
H14.9.6	15405	12143	43
JH17.6.7	14115	11975	32
SK212.4.7	13945	11840	26
JH6.7.7	13715	11756	49
JH6.6.7	13847	11652	32
SK42.7.7	13227	11425	38
PD9.9.6	13950	11421	28
H5.5.7	13667	11251	33
C3.8.6	13642	11137	48
M1.9.6	13818	10993	29
LG2.6.7	12312	10773	37
JH6.7.6	11899	10419	42
H5.8.6	12424	10332	29
C8.E9.6	12085	10251	26
LG2.6.6	11868	10242	36
PD3.8.6	11725	10198	34
M1.7.7	11896	10091	29
PD9.8.6	11703	9713	36
M3.8.6	11245	9520	28
M3.7.7	11338	9496	41
C8.8.6	11244	9438	38
LG2.8.6	11449	9386	33
JH6.8.6	11342	9280	50
JH17.7.7	10153	8818	38
H4.5.7	9840	7911	37
JH17.8.6	9107	7647	43
C2.6.7	8927	7585	26
C3.6.7	8431	7399	29
SK212.7.7	5825	5186	43
Total	**2013001**	**1677331**	
	Mean 20,968.76	Mean 17472.20	Mean 35.54

APPENDIX IX – Relative abundance in log-scale of all 69 microbiota as identified to their lowest taxonomic level but no lower than genus



Final discussion

The first aims of this PhD originated from discussions with beekeepers about the possibilities of widespread breeding for resistance to Varroa in Irish bees. This led to discussion about whether *A. m. mellifera* existed in a pure form in Ireland, in the face of its localised extinction across much of Europe, through replacement with other honey bee sub-species (Meixner *et al.* 2010), and introgression (Jensen *et al.* 2005, Soland-Reckeweg *et al.* 2009). The extent of the latter is particularly pertinent regarding the existence of "wild" colonies which may lose the adaptations that have evolved for their locale (Todesco *et al.* 2016). At the start of this research, it was generally accepted that wild *A. m. mellifera* no longer existed in Ireland. However, contrary to this were direct reports from beekeepers that "wild" colonies existed, that they had been collecting swarms from them for many years and that the bees were consistent with native Black bees, as *A. m. mellifera* is known in Ireland. In addition to the anecdotal evidence, there was tantalising genetic evidence that *A. m. mellifera* existed in Ireland (Jensen *et al.* 2005), including the possibility of a population outside of managed colonies. However, beyond some informal morphometry studies by BKAs and the small amount of genetic evidence from Jensen *et al.* (2005) and Jaffe *et al.* (2010), there was considerable ambiguity about their existence.

The research presented here has now shown that pure *A. m. mellifera* does exist in Ireland, not only in the managed colonies of NIHBS members but also in free-living colonies. Work to find varroa resistant colonies in the managed population was not possible, mainly because the inherent risk to apiaries of non-treatment was too great to tempt beekeepers away from their usual varroa control methods. As a consequence, efforts were redirected towards a different source of putatively resistant colonies in the unmanaged, and therefore untreated, population. As a result of the initial citizen scientist survey, I have shown here that there are a substantial number of free-living colonies nationwide. The survivor colonies of over three years from this study may yet form an important genetic reservoir to assist the managed population in surviving Varroa without the financial cost, and contamination (Calatayud-Vernich *et al.* 2018) and effects on bee health (Johnson *et al.* 2013) of chemical control. The work on free-living bees has now expanded and a collaboration is currently operating with Ireland's National biodiversity data centre which houses a permanent citizens' reporting tool for free-living colonies, "The wild honey bee study" (Moro *et al.* 2021). Further collaboration

with the Survivors taskforce of COLOSS, a honey bee research network, in the creation of "Honey bee watch" aims to make these data of international benefit.

The evidence of pure *A. m. mellifera* produced from the initial mitochondrial and microsatellite data (Hassett *et al.* 2018) generated considerable interest within Ireland's beekeeping community and national media. Whilst it was already recognised that the supply of native Irish queens was insufficient to meet the needs of beekeepers interested in using *A. m. mellifera* over other sub-species or strains, the evidence presented in chapter 1 was a vindication for NIHBS and added an impetus to strengthen their queen breeding scheme. They now intend that each member becomes proficient in queen rearing and passes on their knowledge, and locally raised native queens, to the extended beekeeping community.

The additional evidence of sub-specific purity relative to the European population generated by the SNPs data in my subsequent international collaborations (Henriques *et al.* 2020, Browne *et al.* 2021) combined with engagement with Societas Internationalis pro Conservatione *Apis melliferae melliferae* (SICAMM) and COLOSS has helped move Ireland nearer the spotlight centred on international honey bee research. Furthermore, the outcome of this research has helped move the conversation between NIHBS and the government of Ireland forward towards significant protections for the native ecotype, including the Protection of the native Irish honey bee bill 2021 (Martin *et al*. 2021). This is not only a significant step for Ireland but is likely to be watched with interest across Europe as it has implications for the protection of all local honey bee ecotypes (Fontana *et al.* 2018), a significant block to which has been the European Common Agricultural Policy (CAP) allowing free trade of domestic livestock including honey bees. The tendency in agricultural discussions towards the singular delineation of honey bees as domesticated is undoubtedly an error, particularly wherever it is native. Proximately, honey bees were wild animals (Crane 1984) and, unlike domestic pigs *Sus scrofa domesticus* for instance, where the effects of genetic isolation from their wild conspecific *Sus scrofa* and selective breeding have become evident in their phenotype (Hunter 2018), there is little or no evidence to suggest significant widespread alterations to the honey bee phenotype or genotype in managed honey bee stocks relative to the wild population. Although intensive breeding programmes make it possible to alter phenotypes provided sufficient control on all queen matings is maintained (Cobey 2007), in selective breeding such as for varroa-resistance, it is inadvisable to focus on

only a single desirable characteristic (Uzunov *et al.* 2017) which leaves the door open for genetic variability but also for all ancestral traits. The reality is that the majority of beekeepers do not exercise control on their stocks to the extent required to fix specific characteristics. Despite human interventions which accelerated since the invention of the Langstroth hive, colonies still vary considerably in honey production, defensiveness, propolis production, overwintering capability, and swarming tendency. Irish mitotype data supports this argument as work presented in chapter 2 found no significant differentiation (Fst value 0.013) between the managed and free-living colonies, although there were genotypes unique to each group.

It is probable that an Irish honey bee population originally existed in genetic isolation from mainland Europe and Britain following either natural expansion during the glacial retreat (Carreck 2008), or transport to Ireland during ancient human migration. While the genetic signature of any founder population may be difficult to determine against the background noise of subsequent importations, both historical and modern, the data collected in this study provides hints of mitochondrial genotypes unique to the island. Although the putatively unique genotypes may mean that honey bees in Ireland have differentiated from the European *A. m. mellifera* metapopulation, it would be prudent for additional reference genotypes to be included in future analysis especially from countries outside of those already looked at. If future work permits, it would also be of benefit to attempt the creation of a larger dataset to estimate the divergence time between the European, British and Irish genotypes and paint a clearer picture of their evolutionary dynamics. Certainly, the new links forged by Ireland during this research to researchers in the United Kingdom and mainland Europe could aid in its creation.

Even if future work indicates some genotypes presented here are unique to Ireland, the mitochondrial, microsatellite and SNP genotyping used in this study do not provide evidence of local adaptation. Therefore, carrying on from the work presented here, a genomics approach has started and will be expanded with a particular focus on free-living colonies, as a result of a multi-year research grant having been approved from Science Foundation Ireland. This collaborative work will search the genome of Ireland's bees for functional SNPs known to be associated with characteristics important in resistance to *V. destructor* (Spötter *et al.* 2016) along with other markers associated with metabolic function or muscle development

that may indicate adaptation in the Irish ecotype. It is further intended, using pooled DNA from 30 bees per colony, to analyse >100,000 SNP loci for a more precise assessment of hybridisation and also to look for evidence of natural selection in the presence of selective sweeps (Nielsen *et al.* 2005, Pavlidis and Alachiotis 2017).

The data on pollen add insight into our limited understanding of the foraging activities of honey bees in Ireland which currently consists of only one peer reviewed article (Coffey and Breen 1997) although some evidence exists from honey studies (Downey *et al.* 2005). The Coffey and Breen (1997) data pre-date the discovery of *V. destructor* in the country in 1998 and hence misses any effect that resultant colony losses may have had on the diversity and subsequent foraging activity (Page *et al.* 1995). The depauperate natural flora in Ireland and Britain relative to neighbouring mainland Europe (Grime 1984) combined with the loss of natural habitat harbours the possibly of shifting bee foraging effort more towards a fixed number of crop species. The potential for nutritional stress consequential on such a shift (Naug 2009) is one reason why it is especially important to continue to elucidate the subject of pollen use by all bee species in Ireland using the light of new techniques to accelerate the process. The pollen data for Spring 2017 gave an indication of a shift to crop flora, with a high relative abundance of *Brassica napus* and an unexpectedly low presence of vernal plants. Notably, one colony in the single western apiary showed considerably more pollen diversity than those in the south-east, perhaps reflecting the less intensive crop growing in the west of the country. To expand this work will involve strengthening our collaborative ties with University of Wurzburg who aided us in producing both pollen and bacteria datasets for analysis. Additionally, the extent to which coding skills would be required for in-house analysis of this block of research was underestimated and is something which would require addressing going forward.

It is gratifying to see that this work will be built upon with the new genomics funding, although this is primarily because I have a strong compulsion to reinforce the data on the free-living bees as I feel the picture needs fleshing out. However, the funding to sample and genotype more of these intriguing colonies did not exist at the time. Then there is the question of exactly how *A. m. mellifera* did survive here, particularly in the free-living population where all queens are free-mated. Theories for this survival include assortative mating and sub-

specific stratification or temporal variation in drone congregation areas (Koeniger *et al.* 1989, Rowell *et al.* 1992, Jaffe *et al.* 2009).

The work on this fascinating insect is never-ending however I suspect that no matter what is discovered in the newly funded research, the most gratifying future outcome of this work will be the, I hope inevitable, legal protections for the pure *Apis mellifera mellifera* on the island of Ireland.

References

Browne, K. A., Hassett, J., Geary, M., Moore, E., Henriques, D., Soland-Reckeweg, G., Ferrari, R., Mac Loughlin, E., O'Brien, E., O'Driscoll, S., Young, P., Pinto, M. A. and McCormack, G. P. (2021) 'Investigation of free-living honey bee colonies in Ireland', *Journal of Apicultural Research,* 60(2), 229-240.

Calatayud-Vernich, P., Calatayud, F., Simó, E. and Picó, Y. (2018) 'Pesticide residues in honey bees, pollen and beeswax: Assessing beehive exposure', *Environmental Pollution,* 241, 106-114.

Carreck, N. L. (2008) 'Are honey bees (*Apis mellifera* L.) native to the British Isles?', *Journal of Apicultural Research,* 47(4), 318-322.

Cobey, S. W. (2007) 'Comparison studies of instrumentally inseminated and naturally mated honey bee queens and factors affecting their performance', *Apidologie,* 38(4), 390-410.

Coffey, M. F. and Breen, J. (1997) 'Seasonal variation in pollen and nectar sources of honey bees in Ireland', *Journal of Apicultural Research,* 36(2), 63-76.

Crane, E. (1984) 'Honeybees' *Evolution of domesticated animals*, Ch. 65, pp. 403-415 ed. I.L. Mason. London: Longman Group

Downey, G., Hussey, K., Kelly, J. D., Walshe, T. F. and Martin, P. G. (2005) 'Preliminary contribution to the characterisation of artisanal honey produced on the island of Ireland by palynological and physico-chemical data', *Food Chemistry,* 91(2), 347-354.

Fontana, P., Costa, C., Di Prisco, G., Ruzzier, E., Annoscia, D., Battisti, A., Caoduro, G., Carpana, E., Contessi, A., Dal Lago, A., Dall'Olio, R., De Cristofaro, A., Felicioli, A., Floris, I., Fontanesi, L., Gardi, T., Lodesani, M., Malagnini, V., Manias, L., Manino, A., Marzi, G., Massa, B., Mutinelli, F., Nazzi, F., Pennacchio, F., Porporato, M., Stoppa, G., Tormen, N., Valentini, M. and Segre, A. (2018) 'Appeal for biodiversity protection of native honey bee subspecies of *Apis mellifera* in Italy (San Michele all'Adige declaration)', *Bulletin of Insectology,* 71(2), 257-271.

Grime, J. P. (1984) 'The ecology of species, families and communities of the contemporary British flora', *The New Phytologist,* 98(1), 15-33.

Hassett, J., Browne, K. A., McCormack, G. P., Moore, E., Soland, G., Geary, M. and Native Irish Honey Bee, S. (2018) 'A significant pure population of the dark European honey bee (*Apis mellifera mellifera)* remains in Ireland', *Journal of Apicultural Research,* 57(3), 337-350.

Henriques, D., Lopes, A. R., Ferrari, R., Neves, C. J., Quaresma, A., Browne, K. A., McCormack, G. P. and Pinto, M. A. (2020) 'Can introgression in M-lineage honey bees be detected by abdominal colour patterns?', *Apidologie*.

Hunter, P. (2018) 'The genetics of domestication: Research into the domestication of livestock and companion animals sheds light both on their "evolution" and human history', *EMBO reports,* 19(2), 201-205.

Jaffe, R., Dietemann, V., Allsopp, M. H., Costa, C., Crewe, R. M., Dall'olio, R., de la Rua, P., El-Niweiri, M. A. A., Fries, I., Kezic, N., Meusel, M. S., Paxton, R. J., Shaibi, T., Stolle, E. and Moritz, R. F. A. (2010) 'Estimating the density of honeybee colonies across their natural range to fill the gap in pollinator decline censuses', *Conservation Biology,* 24(2), 583-593.

Jaffe, R., Dietemann, V., Crewe, R. M. and Moritz, R. F. A. (2009) 'Temporal variation in the genetic structure of a drone congregation area: an insight into the population dynamics of wild African honeybees (*Apis mellifera scutellata*)', *Molecular Ecology,* 18(7), 1511-1522.

Jensen, A. B., Palmer, K. A., Boomsma, J. J. and Pedersen, B. V. (2005) 'Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honeybee, *Apis mellifera mellifera*, in northwest Europe', *Molecular Ecology,* 14(1), 93-106.

Johnson, R. M., Dahlgren, L., Siegfried, B. D. and Ellis, M. D. (2013) 'Effect of in-hive miticides on drone honey bee survival and sperm viability', *Journal of Apicultural Research,* 52(2), 88-95.

Koeniger, G., Koeniger, N., Pechhacker, H., Ruttner, F. and Berg, S. (1989) Assortative mating in a mixed population of European honeybees*, Apis mellifera ligustica* and *Apis mellifera carnica*. Insectes Sociaux 36, 129-138

Martin, V.P., McGreehan, E., Garvey, R., Seery-Kearney, M., Boylan, L., Ó Donnghaile, N., Norris, D., Higgins, A-M., Craughwell, G. P., Gavan, P., Warfield, F., Wall, M. (2021), 'Protection of the native Irish honey bee bill 2021' https://www.oireachtas.ie/en/bills/bill/2021/133/, Accessed 04th December 2021

Meixner, M. D., Costa, C., Kryger, P., Hatjina, F., Bouga, M., Ivanova, E. and Buechler, R. (2010) 'Conserving diversity and vitality for honey bee breeding', *Journal of Apicultural Research,* 49(1), 85-92.

Moro, A., Beaurepaire, A., Dall'Olio, R., Rogenstein, S., Blacquière, T., Dahle, B., de Miranda, J. R., Dietemann, V., Locke, B., Licón Luna, R. M., Le Conte, Y. and Neumann, P. (2021) 'Using citizen science to scout honey bee colonies that naturally survive *Varroa destructor* infestations', *Insects,* 12(6).

Naug, D. (2009) 'Nutritional stress due to habitat loss may explain recent honeybee colony collapses', *Biological Conservation,* 142(10), 2369-2372.

Nielsen, R., Williamson, S., Kim, Y., Hubisz, M. J., Clark, A. G. and Bustamante, C. (2005) 'Genomic scans for selective sweeps using SNP data', *Genome research,* 15(11), 1566-1575.

Page, R. E., Waddington, K. D., Hunt, G. J. and Fondrk, M. K. (1995) 'Genetic determinants of honey bee foraging behaviour', *Animal Behaviour,* 50, 1617-1625.

Pavlidis, P. and Alachiotis, N. (2017) 'A survey of methods and tools to detect recent and strong positive selection', *Journal of Biological Research-Thessaloniki,* 24(1), 7.

Rowell, G. A., Taylor, O. R. and Long-Rowell, M. A. (1992) 'Spatial dynamics of the honey bee mating system (*Apis mellifera* L.)', *Journal of the Kansas Entomological Society,* 65(3), 218-222.

Soland-Reckeweg, G., Heckel, G., Neumann, P., Fluri, P. and Excoffier, L. (2009) 'Gene flow in admixed populations and implications for the conservation of the Western honeybee, *Apis mellifera'*, *Journal of Insect Conservation,* 13(3), 317-328.

Spötter, A., Gupta, P., Mayer, M., Reinsch, N. and Bienefeld, K. (2016) 'Genome-wide association study of a varroa-Specific defense behavior in honeybees (*Apis mellifera*)', *The Journal of heredity,* 107(3), 220-227.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hubner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G. and Rieseberg, L. H. (2016) 'Hybridization and extinction', *Evolutionary Applications,* 9(7), 892-908.

Uzunov, A., Brascamp, E. W. and Büchler, R. (2017) 'The basic concept of honey bee breeding programs', *Bee World,* 94(3), 84-87.

Published papers

Journal of Apicultural Research

ISSN: 0021-8839 (Print) 2078-6913 (Online) Journal homepage: http://www.tandfonline.com/loi/tjar20

A significant pure population of the dark European honey bee (Apis mellifera mellifera) remains in Ireland

Jack Hassett, Keith A Browne, Grace P McCormack, Elizabeth Moore, Native Irish Honey Bee Society, Gabrielle Soland & Michael Geary

To cite this article: Jack Hassett, Keith A Browne, Grace P McCormack, Elizabeth Moore, Native Irish Honey Bee Society, Gabrielle Soland & Michael Geary (2018) A significant pure population of the dark European honey bee (Apis mellifera mellifera) remains in Ireland, Journal of Apicultural Research, 57:3, 337-350, DOI: 10.1080/00218839.2018.1433949

To link to this article: https://doi.org/10.1080/00218839.2018.1433949

Journal of Apicultural Research, 2018
Vol. 57, No. 3, 337–350, https://doi.org/10.1080/00218839.2018.1433949

ORIGINAL RESEARCH ARTICLE

A significant pure population of the dark European honey bee (*Apis mellifera mellifera*) remains in Ireland

Jack Hassett[a], Keith A Browne[b], Grace P McCormack[b], Elizabeth Moore[a], Native Irish Honey Bee Society[c], Gabrielle Soland[d] and Michael Geary[a]*

[a]*Department of Applied Science, Limerick Institute of Technology, Limerick, Ireland;* [b]*Zoology, School of Natural Sciences & Ryan Institute, National University of Ireland Galway, Galway, Ireland;* [c]*NIHBS Secretary, Limefield, Clonminch, Tullamore, Co., Offaly, Ireland;* [d]*Apigenix, Twann, Switzerland*

(*Received 14 July 2017; accepted 22 January 2018*)

The natural range of the dark European honey bee, *Apis mellifera mellifera* has been significantly reduced in recent years as a result of importation and replacement of queens with those of other *Apis* subspecies. Previous studies have indicated that a substantial amount of *A. m. mellifera* populations throughout Europe are heavily hybridized but that pockets of pure populations do still exist and need to be protected as this subspecies is a highly valuable gene pool and is of considerable conservational interest. Small numbers of Irish bees have been included in previous studies, but scientific information is limited and questions remain about the genetic diversity of bees in Ireland and the extent of introgression into apparent black bees from introduced races and hybrids. The objective of this study was to investigate the genetic composition of the *A. m. mellifera* population on the island of Ireland with both nuclear (microsatellites) and mitochondrial markers. Molecular data was generated from 412 bees sampled from 24 counties across the island. Mitochondrial data identified 34 different haplotypes, with 63% of bees having sequences identical to three European haplotypes but all other haplotypes being novel. Population structure analysis using microsatellite markers indicates that the Irish population is genetically diverse and that 97.8% of sampled bees were determined to be pure *A. m. mellifera*. Results from cluster analysis using a Bayesian model approach, and the presence of novel alleles, shows evidence of distinctiveness within the Irish population.

Poblaciones significativamente puras de la abeja melífera europea negra (*Apis mellifera mellifera*) permanecen en Irlanda

El área de distribución natural de la abeja melífera europea negra, *Apis mellifera mellifera*, se ha reducido significativamente en los últimos años como resultado de la importación y sustitución de reinas por reinas de otras subespecies de *Apis*. Estudios previos indicaron que una cantidad sustancial de poblaciones de *A. m. mellifera* en toda Europa están fuertemente hibridadas, pero que todavía existen bolsas de poblaciones puras que necesitan ser protegidas, ya que esta subespecie es una reserva genética muy valiosa y de considerable interés conservacionista. En estudios anteriores se incluyó un pequeño número de abejas irlandesas, pero la información científica es limitada y sigue habiendo dudas sobre la diversidad genética de las abejas en Irlanda y el grado de introgresión en las abejas aparentemente negras con razas e híbridos introducidos. El objetivo de este estudio fue investigar la composición genética de la población de *A. m. mellifera* en la isla de Irlanda con marcadores nucleares (microsatélites) y mitocondriales. Se generaron datos moleculares de 412 abejas de 24 condados de la isla. Los datos mitocondriales identificaron 34 haplotipos diferentes, siendo el 63% de las abejas con secuencias idénticas a tres haplotipos europeos, pero todos los demás haplotipos nuevos. El análisis de la estructura de la población utilizando marcadores microsatélites indica que la población irlandesa es genéticamente diversa y se determinó que el 97,8% de las abejas muestreadas eran *A. m. mellifera* pura. Los resultados del análisis de clusters utilizando un enfoque de modelo bayesiano, y la presencia de nuevos alelos, muestran evidencia del carácter distintivo dentro de la población irlandesa.

Keywords: *Apis mellifera*; mitochondrial DNA; microsatellites; conservation; population genetics

Introduction

Following the last glacial period, from approximately 110,000–11,000 years BCE, honey bees (*Apis mellifera*) retreated to the relative safety of the Iberian Peninsula and the Balkan Peninsula. Once the European glaciers had retreated, *A. mellifera* was free to re-colonize Europe with the C Lineages (*A. m. ligustica*, *A. m. carnica*, *A. m. cecropia* and others) in central Europe and the M lineages (including *A. m. mellifera* and *A. m. iberiensis*) becoming established in north and west Europe. Geographical barriers such as the Alps, Pyrenees and Balkan mountains, plus divergent ecological factors, have aided isolation of these lineages leading to the different subspecies and races known today (Han, Wallberg, & Webster, 2012; Hewitt, 1999; Jensen, Palmer, Boomsma, & Pedersen, 2005; Jensen & Pedersen, 2005; Miguel, Iriondo, Garnery, Sheppard, & Estonba, 2007; Miguel et al., 2011; Ruttner, 1988).

*Corresponding author. Email: Michael.Geary@lit.ie

During the northward migration of *A. m. mellifera*, it crossed the land bridge across the Dover Strait into Britain and Ireland (Carreck, 2008; Pritchard, 2009). Following the collapse of the land bridge, the Irish population began what was approximately 6,500 years of isolation. Given the propensity for divergence into different subspecies and races in honey bees, the substantial period of isolation of Irish bees from the rest of mainland Europe could have resulted in the evolution of unique genetic variants. Anecdotally, the 'Irish' bee is described as having a distinct morphology which assists them in surviving in the damp cool climate found on the island. Reports also include peculiarities in foraging and brood rearing behavior in the Irish honey bee. Whether these features are "Irish" or characteristic of all north-western honey bee populations has not been determined.

However human impact on *A. mellifera* has accelerated in the last century to such an extent that the geographic boundaries of sub-species, once regarded as clearly defined, have been radically altered (De la Rúa, Jaffé, Dall'Olio, & Muñoz, 2009). The picture is complicated with gene flow between honey bee sub-species now common within a geographic area due to the importation of non-native subspecies and hybrids and the subsequent difficulty in controlling mating in comparison to other domesticated animals (Franck, Garnery, Solignac, & Cornuet, 2000; Oleksa, Chybicki, Tofilski, & Burczyk, 2011; Oleksa, Wilde, Tofilski, & Chybicki, 2013; Soland-Reckeweg, Heckel, Neumann, Fluri, & Excoffier, 2009). Because of free trade between European countries and thus "artificial" movement of bees according to commercial beekeepers' preferences, a radically different picture of the honey bee population can exist compared to what might be expected via evolutionary forces across natural boundaries and buffer zones. It is likely that a substantial amount of the *A. m. mellifera* population throughout Europe is now heavily hybridized causing concern due to the loss of biological diversity and the possible extinction of the subspecies in past strongholds (Meixner et al., 2010; Pinto et al., 2014; Soland-Reckeweg et al., 2009).

Local ecotypes are reported to be the best bees to use in apiculture due to their adaptation to local conditions, adaptations which if lost cannot be replaced (Parejo et al., 2016; Parker et al., 2010; Szabo & Lefkovitch, 1988). Evidence suggests that the use of local honey bee populations also provides a higher chance of colony survival, and that the use of maladapted bees attributes to high colony losses, as recently observed in many regions (Büchler et al., 2014). Utilizing locally adapted subspecies and ecotypes to buffer populations against various stressors is thus an essential tool in honey bee management (Neumann & Carreck, 2010) and protection of honey bee diversity is therefore crucial as genetic diversity protects the evolutionary potential of species to adapt by natural selection in the future (Allendorf, Luikart, & Aitken, 2012; Frankham, Ballou, Jonathan, & Briscoe, 2010; Mikheyev, Tin, Arora, & Seeley, 2015; Tarpy, 2003).

The influx of non-native bees into the natural range of *A. m. mellifera* has been exacerbated in northern Europe due to long winters and the desire for early queens which can only be provided in more temperate climates. Commercial bee breeding has led to significant imports of exotic subspecies, particularly *A. m. ligustica* from Italy, and *A. m. carnica* from the former Yugoslavia (De la Rúa et al., 2009; Franck et al., 2000; Meixner et al., 2010; Pinto et al., 2014; Ruottinen et al., 2014). As a result, *A. m. mellifera* is reported as the subspecies likely to be under most threat, being virtually replaced in Germany by *A. m. carnica* and largely introgressed by genes from other subspecies over much of the rest of its range (Bouga et al., 2011; Pinto et al., 2014). Thus, much focus has been placed on finding and conserving remaining pure populations of *A. m. mellifera,* and a European initiative to contribute to this includes the establishment of the International Association for the Protection of the European Dark Bee (SICAMM) for a collaborative network and biennial international meeting focusing on the problem of *A. m. mellifera* conservation.

A. m. mellifera is the subspecies of honey bee that is native to Ireland. It has been suggested that *A. m. mellifera* was entirely eliminated from Britain and Ireland at the time of the "Isle of Wight disease" (Adam, 1983). Importation of non-native bees after this population crash is reported to also have had a large impact with the first organized importation of bees recorded in the Republic of Ireland in 1923, when skeps of Dutch bees (*A. m. mellifera*) were brought in large numbers especially to Co. Wexford. In 1927 under a Department of Agriculture and technical instruction restocking scheme, 15 County Committees of Agriculture imported Dutch skeps. The first organized importation of the Italian bee (*A. m. ligustica*) probably occurred in 1927, when 27 queens from the USA and four from England were imported into Wexford (Mac Giolla Coda, 2012). More recently, official data from the Irish Department of Agriculture, Food and the Marine indicate that on average 115 queens have been imported into the Republic of Ireland each year over the last five years. On the island of Ireland, there are a number of "Buckfast" breeders along with some beekeepers who import queens from across Europe on a regular basis.

This history has led to a widespread opinion amongst the public as well as a significant proportion of beekeepers that: (A) Ireland has no native bees; (B) the *A. m. mellifera* population that exists is heavily hybridized; and (C) local breeding groups focused on *A. m. mellifera* results in inbreeding, further affecting the evolutionary potential of this subspecies in Ireland. These views hamper conservation initiatives.

However, recent reassessment of available data on the "Isle of Wight Disease" suggests, that the underlying cause was unlikely to have been solely *Acarapsis woodi*, but rather a combination of mitigating factors including rural population decline during the war. Further, and more importantly, the extent of colony losses was

greatly exaggerated by Brother Adam (Cooper, 1986; Bailey & Ball, 1991; Bailey, 2002; Mac Giolla Coda, 2012). Despite the reports of possible extinction of *A. m. mellifera* in Ireland, and the importation of non-native bees, anecdotal evidence suggests that a considerable pure population of *A. m. mellifera* may exist based on the fact that importation levels have been relatively minor in relation to the expected population size of resident bees and given a substantial effort from key bee breeding organizations, led by the Galtee Bee Breeding Group (GBBG) to protect the black bee in Ireland. In more recent years the Native Irish Honey Bee Society (NIHBS), established in 2012, is developing a conservation program and promoting the conservation of the native dark bee. Preservation of *Apis m. mellifera* is also a constitutional aim of the Federation of Irish Bee Keeping Associations (FIBKA).

A morphometry study by the GBBG on 1040 bees focusing on cubital index and discoidal shift indicated that over 46% of all samples analyzed had a purity greater than 75% with fewer than 5% of samples showing purity less than 25% (Williams, 2013). Only one study investigating genetic diversity of *A. m. mellifera* across Europe included Irish samples (i.e., (Jensen et al., 2005). The authors found just two mitochondrial variants from an Irish sample of 48 bees, which grouped with sequences from the UK and Netherlands. Microsatellite data did not distinguish the Irish population from the British population but the consensus was that the bees examined were consistent with *A. m. mellifera* and had little if any introgression from the C lineage. The samples, although spanning two Irish provinces came from the GBBG and thus may be perceived to be restricted in diversity due to the elements of controlled breeding introduced to protect the black bee in that group of breeders. Protected populations of *A. m. mellifera* from other places in Europe (Denmark, Netherlands, Colonsay Island (Scotland), France, Belgium, Norway and Switzerland) showed introgression from the C lineage subspecies (i.e., *A. m. carnica* and *A. m. ligustica*) by mitochondrial and single nucleotide polymorphism (SNP) data (Pinto et al., 2014). Given the limited scientific information available on Irish bees, questions remain about the genetic diversity of bees in Ireland and the extent of introgression into apparent black bees from introduced races and hybrids. If the anecdotal information is correct, then the Irish population may be important for overall conservation strategy of this subspecies in Europe. Therefore, the objective of the current study is to investigate the genetic composition of the *A. m. mellifera* population on the island of Ireland utilizing both nuclear (microsatellites) and mitochondrial (sequence data from the COI-COII region) markers to help determine any evidence for introgression from non-native honey bee subspecies and hybrids and explore the relationship of Irish bees to *A. m. mellifera* populations in Europe.

Materials and methods

Sampling and DNA extraction

Molecular data was generated from 412 honey bees from 80 sampling sites spread across 24 counties and all four provinces of Ireland (Figure 1 and Online Supplementary Material Table S1). All bees were stored at

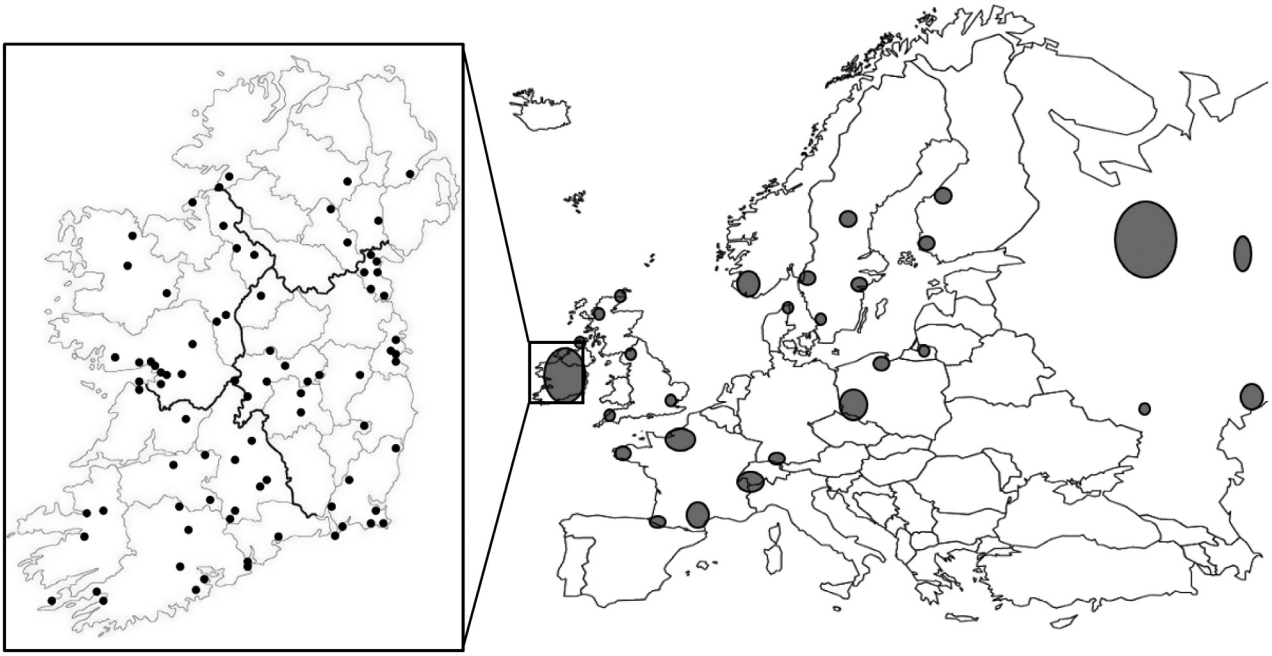


Figure 1. The current distribution areas of *A. m. mellifera* across Europe as highlighted by the shaded areas.
Notes: The enlarged image shows the location of the sampling sites for the analysis of the Irish population.
Source: Jensen et al. (2005), Soland-Reckeweg et al. (2009), Il'yasov et al. (2011) and Il'yasov et al. (2015).

−20 °C before being processed further (Jensen et al., 2005). Where feasible ≥2 bees per colony were included with total DNA being extracted from the hind legs of each bee using the E.Z.N.A DNA extraction kit as per the manufacturers guidelines (Omega bio-tek, 2013).

Mitochondrial DNA

The mtDNA region including the tRNAleu gene, the COI-COII intergenic region and the 5′ end of the COII subunit gene were amplified with the primers E2 and H2 (Garnery, Vautrin, Cornuet, & Solignac, 1991; Meixner et al., 2013) using illustra PuReTaq Ready-To-Go PCR Beads. The reactions were subjected to an initial denaturation of 11 min at 95 °C, followed by 30 cycles of 92 °C for 1 min, 54 °C for 45 sec and 62 °C for 2 min, and a final extension of 10 min at 72 °C (Soland-Reckeweg et al., 2009). PCR products were purified using the GeneJet PCR Purification kit. PCR fragments were sequenced externally by LGC Genomics or Ecogeneics via Sanger sequencing. Resulting chromatographs were assessed manually in MEGA6 before the consensus sequence was imported into a multiple alignment with all other sequences (Evans et al., 2013; Tamura, Stecher, Peterson, Filipski, & Kumar, 2013). A subset of sequences was entered into the BLASTn algorithm in GenBank (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to confirm identity as being of honey bee origin using the "somewhat similar sequences" option due to the large indels present in this region.

A reference alignment containing representative sequences from the most likely subspecies present in Ireland was created using sequences available in GenBank (sequences from *A. m. ligustica, A. m. carnica, A. m. iberiensis, A. m. scutellata*). This alignment was used to screen sequences from Irish bees to determine subspecies status of their mtDNA. Subsequently all those identified as *A. m. mellifera* were aligned to all available European *A. m. mellifera* sequences from GenBank, which were those from Rortais, Arnold, Alburaki, Legout, and Garnery (2011) and Pinto et al. (2014). Unfortunately, sequences from Jensen et al. (2005) were not available. The multiple alignment was generated manually due to the variable numbers of Q elements present in the data-set. Apart from the known Q elements there were few other indels present and manual alignment was straightforward. Identical sequences were identified as were new variants of the region from Irish bees. Only one representative sequence for each haplotype was retained in the alignment. Phylogenetic networks were constructed under statistical parsimony using TCS 1.21 (Posada & Crandall, 2001).

The duplication of the Q elements does generate large numbers of gaps in the alignment and the sequences fall naturally into a number of groups, visible by eye, defined principally by the numbers of Q elements present. Details of the molecular mechanism for duplication of the Q elements are unknown leading to uncertainties as to how they should be treated in phylogenetic analyses, e.g., what weighting they should be afforded. Multiple substitutions and indels are present in the Q elements such that the Q elements are no longer identical across individuals. Furthermore, the duplications are hypothesized to have occurred since the divergence between *A. m. mellifera* (up to four Q elements present) and *A. m. carnica* (one Q element present) and are rare events; not occurring in related bee species or *Drosophila*. Such duplications happening many independent times in *A. m. mellifera* throughout its range in the last 200,000 years since the proposed split with *A. m. carnica* is not parsimonious. Therefore, populations of bees with different numbers of Q elements and substitutions/indels contained within them likely represent evolutionary lineages of bees and represent important synapomorphies.

However, to control for the impact of the Q elements two alignments were employed in reconstructing networks of relationships between the sequences. The first alignment contained all available *A. m. mellifera* sequences and all sites of all Q elements (110 taxa and 1400 sites). The first 23 bp were excluded from analysis due to ambiguities and missing data in some sequences). When all sites from the Q elements were included separate networks were formed containing bees with PQ, PQQ, PQQQ and PQQQQ sequence elements. Eleven European sequences were found not to be connected to any network and may represent under sampled lineages. These individual haplotypes were; HQ337436M4, HQ260378M4c, HQ337442M9, HQ337443M10, HQ337444M10a, HQ337449M16, HQ337457M35, HQ260352M41, HQ260373M58. Also, divergent from the main data-set were HQ260345M18 which was linked to HQ260346M20.

The second alignment excluded the very divergent *A. m. mellifera* European sequences above and also contained only sites in the Q elements that showed variation. Therefore, the second alignment contained 98 taxa and 801 sites. Both alignments are available from the authors on request.

Microsatellites

Twelve microsatellite markers (A007, A28, A29, A43, A76, A273, Ac306, Ap1, Ap33, Ap226, Ap289, B24) were selected based on their informative nature for analysis of *A. m. mellifera* (Alburaki et al., 2013; Estoup, Garnery, Solignac, & Cornuet, 1995; Garnery et al., 1998; Meixner et al., 2013; Soland-Reckeweg et al., 2009) and amplified in two multiplex PCR reactions (Soland-Reckeweg, 2006). The PCRs were performed using the Qiagen multiplex PCR kit and consisted of 10 µl reaction volume containing 0.2 µM of each primer and 3 mM $MgCl_2$. The PCR conditions consisted of an initial denaturation for 15 min at 96 °C, followed by 32 cycles of 94 °C for 30 sec, 60 °C for 1.5 min and 72 °C for 1.5 min, and a final extension of 72 °C for 10 min

(Qiagen, 2010). The PCR products were visualized and sized on an Applied Biosystems 3130 Genetic analyzer using the LIZ 500 size standard, followed by analysis of the data with GeneMapper 5 software.

Reference populations used in the analysis of our population samples included: *A. m. mellifera* Sweden (*n* = 6), *A. m. mellifera* France (n = 24), *A. m. mellifera* Norway (*n* = 18), *A. m. mellifera* Switzerland (*n* = 17), *A. m. ligustica* Italy (*n* = 55), *A. m. carnica* Austria (*n* = 62), *A. m. carnica* Slovenia (*n* = 21), *A. m. carnica* Switzerland (*n* = 91).

Expected heterozygosity for each locus in each population and number of alleles were calculated using ARLEQUIN v3.5.2.2 software. Genetic differentiation between populations, computed using unbiased estimates of pairwise Fst values, was calculated by ARLEQUIN v3.5.2.2 software. Identification of genetically similar groups of individuals was obtained with the software STRUCTURE v2.3.4 (Pritchard, Stephens, & Donnelly, 2000). The results were based on simulations of 100000 burn-in steps and MCMC (Markov Chain Monte Carlo algorithm) iterations. The true number of clusters (K) was estimated using the value for ΔK. The program was run for values of $K = 1–9$, while the most likely number of clusters K was calculated according to Earl and vonHoldt (2012) and Evanno, Regnaut, and Goudet (2005). Microsatellite networks were constructed from allele sharing minimum spanning trees calculated by use of EDENetwork software in accordance with Kivelä, Arnaud-Haond, and Saramäki (2015). Networks consist of nodes representing individual bees which are linked by links or edges which represent their relationships

Table 1. Mitochondrial haplotypes sampled from the Irish population of *A. m. mellifera*.

	Name	Distinction	Number	County
	PQQ			
1	*M4e (Ne)*	Used as base sequence for all comparions below	105	14 counties
2	*M4d (Ne)*	A at 1385	50	≥10 counties
3	*M4a (Fr)*	deletion 126, A at 46, c at 97,	3 (1,2)	Lh
4	16IE146	deletion 28–33, A at 1385	11 (8,9)	Mo,Am,Gy,Ty
5	16IE224	T at 474, A at 1385	10 (1,4+)	Wd
6	16IE105	deletion 84–93 and at 1229, A at 1385	4 (1,2)	Cw
7	16IE145	deletion 126,	5 (3,3)	Gy, Mo
8	16IE313	deletion 460–462, A at 46,	2 (1,1)	Te
9	15IE335	delection 266–275, T at 254, T at 331, A at 328	2 (1,1)	Ck
10	16IE222	A at 46, a at 1385	5 (3,3)	Wd, Ty, Gy
11	*M4f (Ne)*	T at 1258, A at 1385	9 (7,7)	Ls, Ck, Ty, Gy
12	16IE391	A at 46, A at 245	2 (1,1)	Oy
13	16IE356	A at 1236,	1	Gy
14	16IE214	insertion (A) at 206, T at 1258, A at 1385	1	Wd
15	16IE103	deletion 23–32, A at 1385	1	Lk
16	15IE249	T at 108, A at 1385	1	Lh
	PQQQ			
1	15IE324	Used as base sequence for all comparions below	6 (4,4)	Ty, Lk, Gy, Mo
2	16IE375	insertion (A) at 1238, A at 469, CTG at 1224–1226, no deletion at 484–493	6 (1,3)	Gy
3	16IE110	insertion at 348 and 553, A at 469, A at 1385, no deletion at 484–493	4 (3,4)	Gy, Dn, Rn
4	16IE197	A at 46, A at 469, no deletion at 484–493	5 (3,3)	Dn, Ky, Gy
5	16IE159	A at 46, A at 469, T at 118, no deletion at 484–493	4 (2,2)	Gy, Oy
6	16IE130	A at 469, T at 1216, insertion at 1228, no deletion at 484–493	2 (1,1)	Gy
7	16IE124	A at 46, C at 97, A at 460, deletion at 126, no deletion at 484–493	1	Gy
8	16IE077	A at 46, A at 469, T at 1279, no deletion at 484–493	1	Ck
9	16IE217	A at 46, A at 469, deletion 1229, no deletion at 484–493	1	Wd
10	*M4a' (UK)*	A at 469, T at 1216, no deletion at 484–493	1	Oy
11	16IE386	A at 469, T at 1344, A at 1385, insertion at 348 and 553, no deletion at 484–493	1	Rn
12	16IE204	A at 1236, no deletion at 484–493	2 (2,2)	Cw, Ky
	PQQQQ			
1	16IE306	Used as base sequence for all comparions below	4 (1,2)	Am
2	16IE363	A at 46, A at 666, A at 1367, G at 1386	1	Gy
3	16IE380	A at 46, A at 862, A at 1382	1	Gy
4	16IE361	A at 46, A at 862, A at 1367	1	Gy
5	16IE158	T at 833, C at 834, A at 1367, G at 1378	1	So
6	15IE175	A at 46, T at 319, insertion of T at 348, A at 666	1	
34			255	

Notes: Name given corresponds to the sample/accession number of the haplotype (sequence of which has been lodged in GenBank or to an already known haplotype if one has been found). Also included are details of substitutions or indels unique to haplotypes, the number of times the haplotype occurred in the Irish population sampled, the numbers of colonies/apiaries in question and the numbers of counties in Ireland that the haplotype was present in. Haplotypes have been divided into PQQ, PQQQ and PQQQQ variant types. Number refers to firstly number of sequences of that haplotype detected (number of apiaries, number of colonies). Lh = Louth, Am = Antrim, Gy = Galway, Ty = Tipperary, Wd = Wexford, Te = Tyrone, Mo = Mayo, Ck = Cork, Ls = Laois, Oy = Offaly, Lk = Limerick, Dn = Dublin, Rn = Roscommon, Ky = Kerry, Cw = Carlow, So = Sligo. Ne = Netherlands, Fr = France, UK = United Kingdom.

and the strength of such relationship or genetic distance is associated with edge width.

Results

Population structure based on mitochondrial data

In total 255 mtDNA sequences were deemed to be of sufficient quality for inclusion. Seven bees, sampled from a total of four colonies, were diagnosed as being of the C lineage as the sequences were missing the P region. Two putative "Buckfast" bees from one colony had sequences that were distinct from both the C lineage sequences and all other Irish *A. m. mellifera* sequences. They both contained the same insertion in the P sequence that is present in M10, M10a, M12, M13 and M16 from Rortais et al. (2011).

Of the sequences that were characterized as being of the M lineage (*A. m. mellifera*) a total of 34 different mitochondrial haplotypes were found amongst Irish bees. There were no sequences of the PQ type present amongst the Irish sequences. There were 16 variants with PQQ pattern, 12 with PQQQ pattern and 6 with a PQQQQ pattern. Table 1 contains details of these haplotypes and their frequencies in the Irish bees sampled. The most common haplotype in the total population was M4e at 41% followed by haplotype M4d at 20% and differing by only one substitution. Both haplotypes were widespread throughout the country being present in 10–14 counties sampled (Table 1).

In total, five European haplotypes were found in the Irish population including M4d and M4e above, M4a, M4f and M4a′. The M4d, M4e and M4f sequences were identical to sequences from the Netherlands and lodged in GenBank (KF274627, KF274628 and KF274629) by Pinto et al. (2014), the M4a sequence was that from France (KF274625) and present in three bees from two colonies from county Louth while the M4a' sequence was from Scotland (KF274638) from the same study and present in one bee from Co. Offaly. The most common uniquely Irish haplotype was found in 11 bees (5%) across nine colonies from four different counties in the south, west and north of Ireland. All remaining haplotypes were present in less than 5% of bees sampled and are listed along with how they were defined and the counties in which they were sampled in Table 1.

When all information from the Q elements were included for analysis, separate networks were generated for PQ, PQQ and PQQQ and PQQQQ variants. However, when divergent haplotypes and large sections of the Q elements were removed from analysis these variants largely still represented separate lineages. Bees with sequences only containing one Q element form a small network as there are few representatives of these bees sampled and none from Ireland (Figure 2).

The largest number of sequences and haplotypes are of the PQQ type. The root of the network was proposed by the TCS program to be a sequence from the Netherlands which is identical to 50 (20%) of the Irish sequences sampled (i.e., KF274627M4d from Pinto et al. (2014). One mutational step away is a haplotype identical to that found in the Netherlands (KF274628M4e) and to 105 (40%) of the sequences sampled from Ireland. A haplotype from France (KF274625M4a, also from Pinto et al. (2014) is at the center of a sub cluster of primarily European sequences though also including 16IE313 from Northern Ireland.

Another network is entirely comprised of PQQQ and PQQQQ sequences (Figure 2) and within this the PQQQQ sequences are largely distinct from the PQQQ sequences, being linked to the proposed root of the network and to one other sequence but otherwise at the edge of the network. Four PQQQ sequences were not connected to this network due to significant indels (15IE324 and 16IE204, & M55 and M9 from Rortais et al., 2011). The haplotype found in Ireland and in Scotland was proposed to be the out group for the PQQQ & PQQQQ lineages (16IE130 – Ireland, this study) and KF274638M4a (Scotland, Pinto et al., 2014). As indicated in Figure 2 all the other PQQQ haplotypes from Ireland are positioned closer to the proposed root than a group of European sequences mostly from France and Belgium sequenced by Rortais et al. (2011). 15IE324 and 16IE304 are >10 mutational steps from the root and some of the European sequences form a separate cluster.

Population structure based on microsatellite markers

The number of alleles present at each locus varied considerably in the Irish population, with some of the loci being highly variable (A76 = 37 alleles) and others showing a lack of variability (A273 = 5) (Table 2).

The average number of alleles per locus within the Irish population was 14.7 compared to 9.3 for the European *A. m. mellifera* populations and 18.9 for the C lineage populations. The high level of alleles per locus in the C lineage samples can be explained by the occurrence of two subspecies within the population sampled, *A. m. carnica* and *A. m. ligustica*. The average genetic diversity, measured as expected heterozygosity, in the Irish population was calculated as 0.558 which is comparable to the expected heterozygosity value of 0.544 calculated for the reference *A. m. mellifera* population (from multiple countries). There was no significant difference between the expected and observed heterozygosity of the Irish population when all loci were analyzed together ($p = 0.58$). The numbers of alleles present in the Irish population for locus Ap 226 was significantly higher than in the European reference data-set (Table 2). However, most individuals sampled were homozygous at this locus involving 2 of the 10 alleles identified

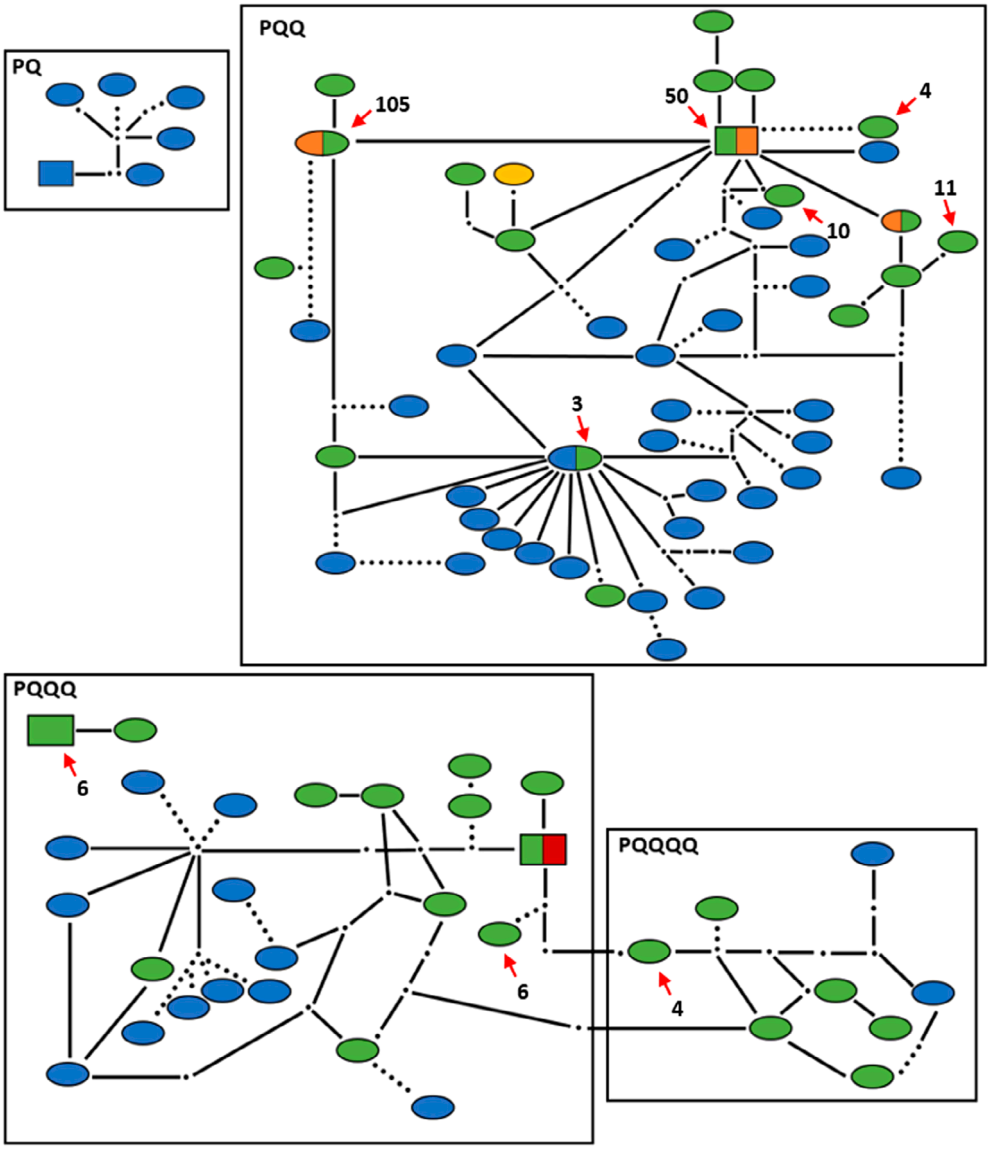


Figure 2. TCS Networks drawn from mitochondrial haplotype sequence data.
Notes: Haplotypes are colored according to geographical location; Green = Irish, Blue = unknown European location (mostly France and Belgium), Red = UK, Orange = Netherlands, Yellow = '*A m. ligustica* type' from Pinto et al. (2014) that groups well inside the *A. m. mellifera* data-set. The numbers beside the boxes refer to the number of sequences of this haplotype found in the Irish bees sampled for this study. Where there are no numbers this means only one sequence was present. Where two colors are side by side, this means that the Irish sequences obtained were identical to an already described haplotype; all other green haplotypes are currently unique to Irish bees. The number of dots represents deletions between connected samples, with the lines representing substitutions.

thus leading to a significantly lower observed heterozygosity than expected. Across each of the loci examined a number of alleles were found to be present within the Irish population which were not found in the European *A. m. mellifera* reference population (Online Supplementary Material Table S2).

Pairwise *F*st values indicate some structure in the data-set with separation indicated between *A. m. mellifera* populations (both Irish and European) and C lineages (*A. m. carnica and A. m. ligustica*) with an *F*st value of >0.4 (Table 3a). A small level of structuring between Irish and European *A. m. mellifera* populations is indicated with an Fst value of 0.072. When focus is placed at investigating structuring between populations from Ireland and different countries in Europe the lowest *F*st value at 0.07 was obtained for between Ireland and France with *F*st values between Ireland and both Norway and Switzerland having values close to 0.1 (Table 3b).

When the Irish population was separated into subpopulations based around centers of significant breeding focus on the black bee, and *F* statistics were used to explore geneflow between them and European populations, the lowest Fst values were between the GBBG bees and bee breeders in Louth and Connemara respectively at 0.04 and 0.05 (Table 4). Fst values between

Table 2. Microsatellite diversity in the Irish population compared with reference populations of *A. m. mellifera* and European C lineages.

Locus	Irish	EU mellifera	EU C lineage	Locus	Irish	EU mellifera	EU C lineage
Ap273				**A76**			
N	317	65	229	*N*	313	66	168
n	5	3	4	*n*	37	21	95
He	0.25161	0.26462	0.24392	He	0.92191	0.91156	0.97877
Ho	0.23344	0.21538	0.23581	Ho	0.86262	0.88889	0.83333
A43				**A007**			
N	317	65	228	*N*	315	65	229
n	6	5	8	*n*	11	5	20
He	0.29556	0.47124	0.67038	He	0.47709	0.20024	0.80391
Ho	0.28076	0.38462	0.51754	Ho	0.48571	0.20000	0.77293
Ac306				**Ap1**			
N	178	65	228	*N*	229	56	171
n	6	5	11	*n*	33	18	20
He	0.58828	0.51030	0.34438	He	0.87749	0.96786	0.55900
Ho	0.48315	0.49231	0.31579	Ho	0.82609	0.86207	0.57310
Ap33				**A28**			
N	313	64	228	*N*	316	65	228
n	16	14	15	*n*	5	4	12
He	0.84320	0.88437	0.85465	He	0.06776	0.209822	0.51893
Ho	0.76358	0.50000	0.75439	Ho	0.05696	0.230769	0.22368
B24				**Ap289**			
N	317	65	229	*N*	315	64	229
n	5	3	5	*n*	19	15	25
He	0.19756	0.25953	0.59874	He	0.88067	0.89330	0.56910
Ho	0.16719	0.26154	0.58079	Ho	0.84444	0.82813	0.27074
Ap226				**A29**			
N	315	63	226	*N*	216	84	174
n	10	3	7	*n*	23	15	25
He	0.43464	0.04710	0.62733	He	0.85615	0.90760	0.91388
Ho	0.06667	0.04762	0.61947	Ho	0.75316	0.83721	0.59770

Notes: *N* = sample size, *n* = number of alleles detected, He and Ho = expected and observed heterozygosity.

Table 3a. Divergence matrix for the microsatellite DNA analysis of the Irish population, the C lineage reference population and the European *A. m. mellifera* reference population.

	Irish	C Lineage	EU mellifera
Irish	0.00000		
C Lineage	0.44777	0.00000	
Eu mellifera	0.07239	0.42487	0.00000

Connemara and Louth (0.0883) were equivalent to the Fst values between GBBG bees and the French *A. m. mellifera* (0.0878) and slightly greater than between bees from Connemara and France (0.0729). Pairwise Fst between the Irish subpopulations and populations from Switzerland and Norway were all >0.1. The Louth Breeding Group bees however showed a closer relationship to the Norway samples than the French

Table 3b. Divergence matrix for the microsatellite DNA analysis of the Irish population, the C reference lineage populations (Italy, Austria, Slovenia and Swiss carnica) and the European *A. m. mellifera* reference population (Sweden, France, Norway and Swiss mellifera).

	Ireland	Italy	Austria	Slovenia	Swiss carnica	Sweden	France	Norway	Swiss mellifera
Ireland	0.00000								
Italy	0.54630	0.00000							
Austria	0.49298	0.38101	0.00000						
Slovenia	0.55183	0.42570	0.06772	0.00000					
Swiss carnica	0.49715	0.38067	0.03904	0.04221	0.00000				
Sweden	0.21847	0.55627	0.48337	0.59515	0.49099	0.00000			
France	0.07412	0.57152	0.48933	0.60187	0.49481	0.24440	0.00000		
Norway	0.10379	0.55116	0.48750	0.59070	0.49649	0.18725	0.10968	0.00000	
Swiss mellifera	0.10621	0.55892	0.46955	0.58134	0.47778	0.24185	0.06331	0.08881	0.00000

Table 4. Divergence matrix for the microsatellite DNA analysis of the Irish Breeding groups (Galtee, Connemara and Louth) and the European *A. m. mellifera* reference population (Sweden, France, Norway and Swiss mellifera).

	Galtee	Connemara	Louth	Sewden	France	Norway	Switzerland
Galtee	0.00000						
Connemara	0.05741	0.00000					
Louth	0.04060	0.08838	0.00000				
Sweden	0.25421	0.21578	0.23591	0.00000			
France	0.08781	0.07293	0.13191	0.24440	0.00000		
Norway	0.12270	0.11266	0.10709	0.18725	0.10968	0.00000	
Switzerland	0.12314	0.10955	0.15387	0.24185	0.06331	0.08881	0.00000

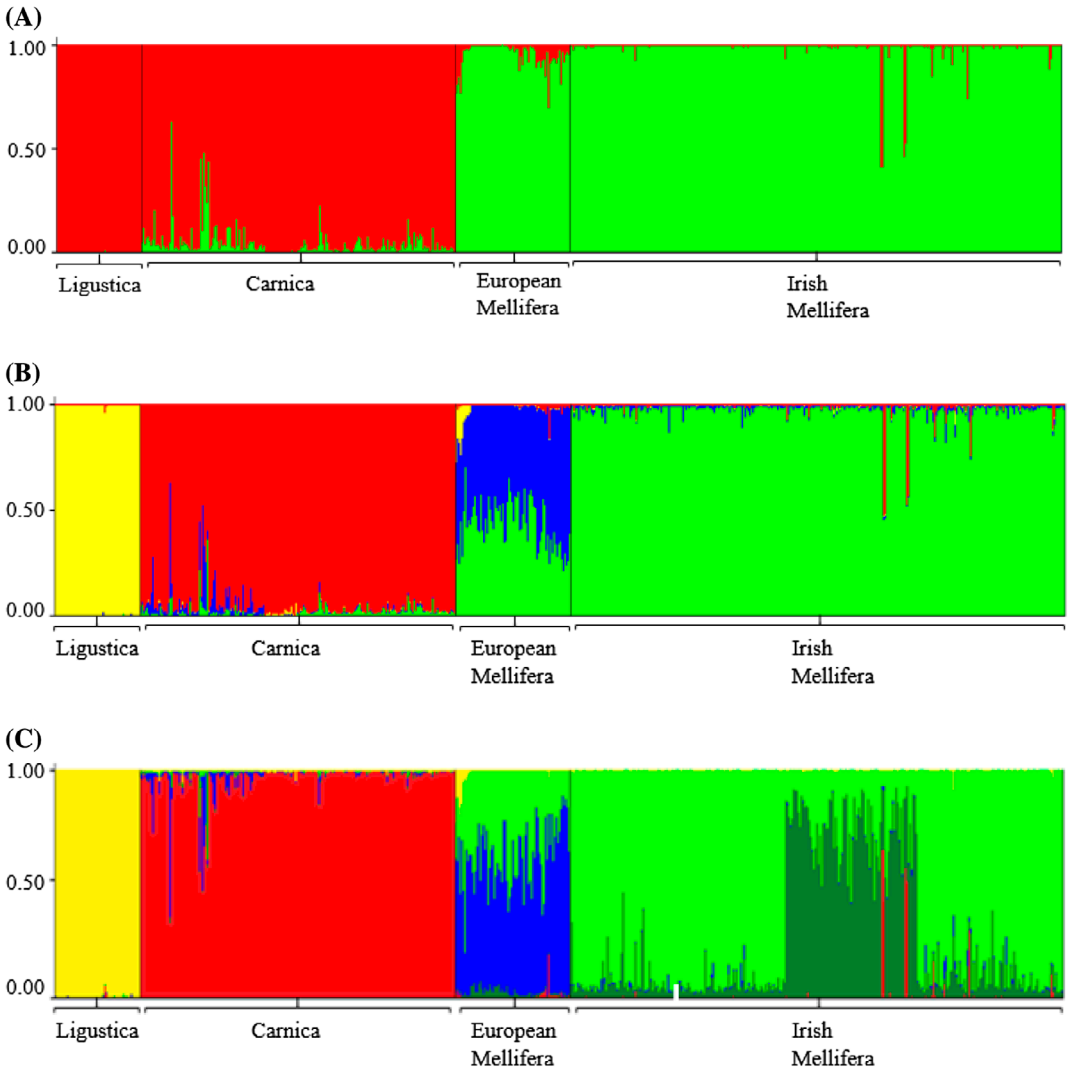


Figure 3. Structure analysis on microsatellite data from Irish and European bees specifying $K = 2$, $K = 4$ and $K = 5$. (A) When $K = 2$, differentiation is evident between the M lineage (green) and the C lineage (red). (B) When $K = 4$, differentiation is evident between the *A. m. ligustica* C population (yellow), *A. m. carnica* C populations (red), the European *A. m. mellifera* M (blue) and the Irish M (green). (C) When $K = 5$, additional differentiation is indicated between two groups of the Irish M population (light green and dark green).

samples even though the *Fst* value was still >0.1 (Table 4).

Population structure and admixture inferred with STRUCTURE for a number of *k* clusters including 2, 4 and 5, are shown in Figure 3. The highest value of ΔK was detected when a model assuming two populations was set but minor peaks were also detected at $K = 4$ and $K = 5$ (Online Supplementary Material Figure S1). At a $K = 2$, each of the main lineages (M and C) grouped clearly into two distinct clusters. The majority of the

Irish bees were grouped into the M lineage along with the European *A. m. mellifera* bees with proportional membership of the Irish bees to this group being 0.992 ($\alpha < 0.05$). For comparison, the proportional membership of the European bees to this group was lower at 0.959 ($\alpha < 0.05$) indicating greater levels of introgression from non-M subtypes in the European bees. On an individual bee level, the lowest calculated proportional membership of any single Irish bee to the M cluster belonged to four bees. Two of these were identified as "Buckfast" honey bees and were included for reference. These samples showed proportional membership to *A. m. mellifera* at 0.393 (16IE315) and 0.399 (16IE316) while two bees sampled from Mayo showed values of 0.443 (16IE144) and 0.514 (16IE145). These approximate half purity values correspond to hybrid status. The "Buckfast" bees sampled in Donegal are hybrid between *A. m. mellifera* and *A. m. carnica*. Interestingly, when a thresholded network analysis was conducted (data not shown), the "Buckfast" samples grouped completely outside the remainder of the data-set, indicating their distinctiveness and hybrid nature. Aside from these hybrids, 97.8% of Irish samples scored higher than 0.96, i.e., higher than the 0.900 necessary to be deemed as pure *A. m. mellifera* bees (Vähä & Primmer, 2006).

Increasing the K cluster number to 4 corresponds to the C lineage being divided into *A. m. carnica* versus *A. m. ligustica* and the M lineage divided into Irish versus European bees (Figure 3) providing evidence of distinctiveness between Irish black bees and European black bees. The European *A. m. mellifera* bees contained a mixture of 'Irish' and 'European' alleles, with an average proportional membership of 0.595 to EU *mellifera* and 0.360 to "Irish" *mellifera*. Some European bees showed a proportional membership of 0.757 to "Irish" *mellifera*. However, the Irish bees show little introgression of European alleles with an average proportional membership of 0.007.

At $K = 5$ there is some indication of sub-clustering within the Irish population (Figure 3). One cluster comprised the majority of the bees sampled (238) and the second cluster comprised 78 bees. This second cluster showed a certain amount of introgression from the first Irish cluster, the proportional membership averaging at approx. 0.251 but this cluster was completely void of introgression from the European *A. m. mellifera* cluster.

A network drawn from the distribution of alleles between Irish and European bees indicate a clear separation of bees of M vs. C lineages. Within the M area of the network (Figure 4) two dense clusters are evident

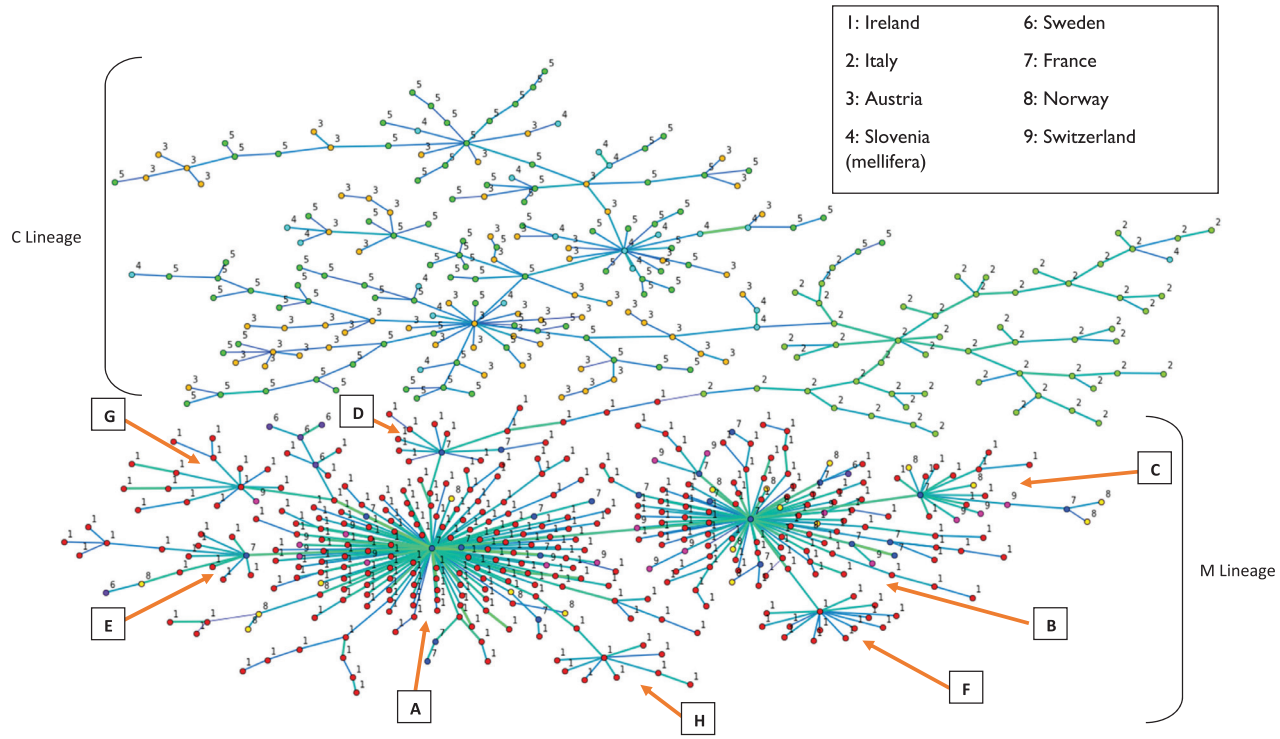


Figure 4. Allele sharing network constructed using EDENetworks detailing the cluster formation (A–H) of the Irish samples and the reference M lineage and C lineage populations.
Notes: Nodes indicate individual bee samples with number attached detailing the location of the sample. The denser the line the closer the relationship between connected nodes.

with French bees being the central node of both but both also containing large numbers of Irish bees. Six smaller clusters were observed with bees from France forming the central nodes in three of them (C, D and E) and containing data from bees from an Irish and European locations. Three clusters however, had Irish bees as the central nodes. Furthermore, two of the clusters contained data only from Irish bees, i.e. 16IE212 (cluster F, 14 Irish bees), 15IE302 (cluster G, 14 Irish plus one Swiss bee) and 15IE281 (cluster H, 10 Irish bees).

When network clustering analysis was performed on just the Irish data-set (Online Supplementary Material Figure S2) two main clusters were evident with 15IE302 (I) (Co. Antrim) and 16IE212 (J) (Co. Wexford) being the central nodes of the clusters with 42 and 32 linked nodes respectively. A smaller sub cluster (K) with 11 bees linked to 15IE281 (Co. Louth) is evident connected to each of the two larger clusters by a number of intermediates. While two main clusters stand out many other Irish variants are also evident in the data showing a significant diversity within the population sampled. Data retrieved from bees sampled from the main breeding centers (Galtee, Louth and Connemara) were intermixed on the network.

Discussion

This study presents the first comprehensive genetic analysis of the indigenous population of *A. m. mellifera* on the island of Ireland. Data from 412 honey bees from 80 sampling sites across 24 counties shows the existence of a highly pure breeding population throughout the island. This is evident in that 97.8% of Irish samples were assigned to the subspecies with a probability of 0.96, when a probability of 0.90 indicates purity. While historical importations of sub-species other than *A. m. mellifera* have been recorded since the 1920s (Watson, 1981) and current legislation also allows for importation, there appears to have been extremely low levels of introgression into the naturalized population sampled here. The possible reasons for this low level of introgression include the island isolation, the comparatively non-commercial nature of beekeeping in Ireland with relatively little importation of non-native subspecies, assortative mating, better survivability under Irish conditions and finally, a preference for localized breeding programs.

Another key finding of this study was the diverse nature of the island population. There was no significant difference between the expected and observed heterozygosity of the Irish population indicating an absence of inbreeding. Furthermore, the expected heterozygosity of the Irish population was greater than that of the C lineages (containing two subspecies though low numbers) and also of the European populations combined (low numbers of individuals but from four countries). While Jensen et al. (2005) showed just two mitochondrial lineages from the 48 Irish bees included and an indistinguishable microsatellite profile compared to UK bees, here we show clearly that this apparent limited genetic diversity was probably a sampling effect. By increasing numbers of bees and sampling sites we report the existence of a significant number of unique alleles and mitochondrial haplotypes, some of which we accept may be found within unsampled European populations in the future.

Evidence from both data types also supports substructures within *A. m. mellifera* as the Irish *mellifera* population was reasonably distinct from the European *mellifera*. A high level of Irish alleles being present in the European reference populations may reflect the common ancestry of these two populations. The very limited gene flow from the European populations sampled into the Irish population indicates the isolation of the Irish population while the presence of unique microsatellite alleles and mitochondrial haplotypes in the Irish population probably indicates independent evolution of the Irish population since its isolation from mainland Europe. Further structuring was observed ($K = 5$) in the Irish *A. m. mellifera* population pointing to a cohort of the Irish population that may be segregated in part from the main population. This cohort showed no evidence whatsoever of introgression from European alleles using microsatellites, highlighting the isolation from the rest of the European population and perhaps hinting at the existence of a native Irish bee that has evolved in separation from other populations in Europe since the closure of the land bridge with Britain.

Some of the historical linkages and the more recent importation influences of the European *mellifera* have been shown in this study. The majority of Irish mitochondrial sequences were identical to three haplotypes that were described from the Netherlands, while one was identical to a French haplotype and another to one from Colonsay Island in Scotland (Pinto et al., 2014). Of the two lineages found by Jensen et al. (2005) from Irish bees one was related to bees from the UK and the other related to bees from the Netherlands. The bees from Jensen et al. (2005) were sampled from GBBG and corroborated the records maintained by the group in that they derived from the population existing in Ireland in the 1920s as well as those imported from the Netherlands after the Isle of Wight disease (Mac Giolla Coda, personal communication, 2016). So, the large numbers of bees similar to the Dutch type detected here reflect the significant imports by beekeepers from the Netherlands after the loss of managed colonies during Isle of Wight disease. Here mitochondrial data coincides with evidence described in the grey literature and by word of mouth. Whether the same levels of Dutch haplotypes will be present in a wider sample of non-managed colonies remains to be seen. But certainly, amongst the beekeepers in the NIHBS these are the predominant type of *A. m. mellifera* here. Indeed, the GBBG sent bees to Colonsay and other locations in the

UK indicating the potential of these types of analyses to detect relationships.

Networks drawn from microsatellite data showed that Irish bees appeared more closely related to French bees than those from Norway or Sweden. For example, two clusters containing primarily Irish bees each had a French bee at its center, perhaps indicating the movement of bees from France through Britain and into Ireland, or more recent French imports of bees to Ireland. The relationship with Dutch and UK bees could not be addressed using microsatellite analysis as regrettably no bees from the Netherlands or the UK were included here and no data were available for use. However, mitochondrial networks also show a relationship between Irish and French bees.

Smaller exclusively Irish sub-clusters were identified in both microsatellite and mitochondrial networks, which may indicate locally adapted ecotypes within the Irish population. Substructures or ecotypes within *A. m. mellifera* have been previously described, e.g., between Swiss and French populations even though there was a relatively short ≈40 km between the conservation area in the French Alps and the Swiss *A. m. mellifera* population. This was possible due to genetic drift, differential bee keeping or local adaptations (Parejo et al., 2016). Within the French population, adaption to local flora and the effects of 'genotype by environment interactions' have resulted in locally adapted honey bee populations which have developed their own resistance mechanisms matching their environment (Meixner, Kryger, & Costa, 2015; Strange, Garnery, & Sheppard, 2008; Tarpy, 2003). However, the Irish bees did not separate according to geography or local breeding groups suggesting some other factors as a cause of any substructuring.

The sub-division of the Irish populations into significant breeding groups also highlights a number of other possible connections and relationships between regions. Low *F*st values confirm some geneflow between Ireland and France. It is however interesting to note that the Louth region appeared to have higher geneflow with Norway rather than France. This may reflect the position of Louth close to the border between Northern Ireland and the Irish Republic and a possible movement of bees from north to south. The source of bees being imported to Northern Ireland (part of the UK) may be quite different from those imported to the Republic of Ireland. While the bees in the Louth breeding group have been sourced primarily from local swarms, the addition of queens from the GBBG can be detected here given that that the Louth and Galtee breeding groups appeared as being more closely associated. Lower levels of geneflow are evident between Connemara and Louth reflecting perhaps the preferred use of local bees by the Connemara breeding group but also a possible reduction in geneflow between east and west.

In conclusion, based on both the mitochondrial and microsatellite results, Ireland is home to a significant pure population of *A. m. mellifera*. This population is comprised of bees that show clear linkages with European bees, particularly from France the Netherlands and the UK, and another group of bees that show distinct "Irish" microsatellite alleles and mitochondrial haplotypes. Together this data indicates a diverse population that does not suffer from inbreeding nor does it suffer from introgression from C lineages despite the continued imports and breeding of non Irish bees in Ireland. The presence of such a widespread and pure population of *A. m. mellifera* in Ireland is now an incredibly important resource for the protection of this subspecies in Europe. Given the devastating impacts of varroa from introduced bees on this population, particularly on wild bees, efforts should be increased to prevent any impact on this subspecies from introductions of pests and diseases due to the continued imports of non Irish bees.

Supplementary material

Supplementary material is available for this article at: https://doi.org/10.1080/00218839.2018.1433949.

Acknowledgements

The Authors would like to thank the beekeepers who kindly provided bees throughout the course of this study. We would like to thank the Mac Giolla Coda family and the Galtee Bee Breeding Group for their invaluable assistance. The Authors are grateful to Limerick Institute of Technology and the National University of Ireland Galway for the use of their facilities.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

JH is funded by the Institutes of Technology of Ireland, KB is a recipient of an Irish Research Council postgraduate fellowship and a Tony Ryan Postgraduate fellowship. Additional funding was gratefully received from the Department of Agriculture, Food and the Marine [grant number GRGAS 16/GR/09], the Federation of Irish Beekeeping Associations, the Eva Crane Trust [grant number ECTA20160303] and The Native Irish Honey Bee Society.

ORCID

Keith A Browne http://orcid.org/0000-0001-6666-9821
Grace P McCormack http://orcid.org/0000-0003-4960-5185

References

Adam, B. (1983). *In search of the best strains of bees and the results of the evaluations of the crosses and races*. Hebden Bridg: Northern Bee Books.

Alburaki, M., Bertrand, B., Legout, H., Moulin, S., Alburaki, A., Sheppard, W. S., & Garnery, L. (2013). A fifth major genetic group among honey bees revealed in Syria. *BMC Genetics, 14*, 1–10.

Allendorf, F.W., Luikart, G., & Aitken, S.N. (2012). *Conservation and the genetics of populations*. Hoboken: John Wiley & Sons.

Bailey, L. (2002). *The Isle of Wight disease*. Parkstone Poole Dorset: Central Association of Bee-Keepers.

Bailey, L., & Ball, B. V. (1991). *Honey bee pathology*. London: Academic Press.

Bouga, M., Alaux, C., Bienkowska, M., Büchler, R., Carreck, N.L., Cauia, E., … Gregorc, A. (2011). A review of methods for discrimination of honey bee populations as applied to European beekeeping. *Journal of Apicultural Research, 50* (1), 51–84. doi: 10.3896/IBRA.1.50.1.06

Büchler, R., Costa, C., Hatjina, F., Andonov, S., Meixner, M.D., Conte, Y.L., … Drazic, M. (2014). The influence of genetic origin and its interaction with environmental effects on the survival of *Apis mellifera* L. colonies in Europe. *Journal of Apicultural Research, 53*(2), 205–214. doi:10.3896/IBRA.1.53.2.03

Carreck, N.L. (2008). Are honey bees (*Apis mellifera* L.) native to the British Isles? *Journal of Apicultural Research, 47*(4), 318–322. doi:10.1080/00218839.2008.11101482

Cooper, B. (1986). *The honey bees of the British Isles*. Derby: Bee Improvement and Bee breeders Association.

De la Rúa, P., Jaffé, R., Dall'Olio, R., Muñoz, I., & Serrano, J. (2009). Biodiversity, conservation and current threats to European honey bees. *Apidologie, 40*, 263–284.

Earl, D.A., & vonHoldt, B.M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*(2), 359–361.

Estoup, A., Garnery, L., Solignac, M., & Cornuet, J.M. (1995). Microsatellite variation in honey bee (*Apis mellifera* L.) populations: Hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics, 140*(2), 679–695.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology, 14*(8), 2611–2620.

Evans, J.D., Schwarz, R.S., Chen, Y.-P., Budge, G., Cornman, R.S., De La Rua, P., … Pinto, M.A. (2013). Standard methodologies for molecular research in *Apis mellifera*. In V. Dietemann, J. D. Ellis & P. Neumann (Eds.), The COLOSS BEEBOOK, Volume I: standard methods for *Apis mellifera* research. *Journal of Apicultural Research, 52*(4), 1–54. doi:10.3896/IBRA.1.52.4.11

Franck, P., Garnery, L., Solignac, M., & Cornuet, J.M. (2000). Molecular confirmation of a fourth lineage in honey bees from the Near East. *Apidologie, 31*, 167–180.

Frankham, R., Ballou, J.D., Jonathan, D., & Briscoe, D.A. (2010). *Introduction to conservation genetics*. Cambridge: Cambridge University Press.

Garnery, L., Franck, P., Baudry, E., Vautrin, D., Cornuet, J.M., & Solignac, M. (1998). Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). 1. Mitochondrial DNA. *Genetics Selection Evolution, 30*(Suppl 1), S31–47.

Garnery, L., Vautrin, D., Cornuet, J.M., & Solignac, M. (1991). Phylogenetic relationships in the genus *Apis* inferred from mitochondrial DNA sequence data. *Apidologie, 22*(1), 87–92.

Mac Giolla Coda, E. (2012). Isle of Wight disease in Ireland. *The Four Seasons, 46*, 13–15.

Han, F., Wallberg, A., & Webster, M.T. (2012). From where did the western honey bee (*Apis mellifera*) originate? *Ecology and Evolution, 2*(8), 1949–1957.

Hewitt, G.M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society, 68*, 87–112.

Il'yasov, R. A., Kutuev, I. A., Petukhov, A. V., Poskryakov, A. V., & Nikolenko, A. G. (2011). Phylogenetic Relationships of Dark European Honey bees Apis mellifera mellifera L. from th Russian Ural and West European Populations. *Journal of Apicultural Science, 55*(1), 67–77.

Il'yasov, R. A., Poskryakov, A. V., Petukhov, A. V., & Nikolenko, A. G. (2015). Genetic differentiation of local populations of the dark European bee *Apis mellifera* mellifera L. in the Urals. *Russian Journal of Genetics, 51*(7), 677–682.

Jensen, A. B., Palmer, K.A., Boomsma, J.J., & Pedersen, B.V. (2005). Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honey bee, *Apis mellifera mellifera*, in northwest Europe. *Molecular Ecology, 14*(1), 93–106.

Jensen, A.B., & Pedersen, B.V. (2005). Honey bee conservation: A case study from Læsø Island (DK). *Beekeeping and conserving biodiversity of honey bees* (pp. 142–164). Hebden Bridge: Northern Bee Books.

Kivelä, M., Arnaud-Haond, S., & Saramäki, J. (2015). EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics. *Molecular Ecology Resources, 15*(1), 117–122.

Meixner, M.D., Costa, C., Kryger, P., Hatjina, F., Bouga, M., Ivanova, E., & Büchler, R. (2010). Conserving diversity and vitality for honey bee breeding. *Journal of Apicultural Research, 49*(1), 85–92. doi:10.3896/IBRA.1.49.1.12

Meixner, M.D., Kryger, P., & Costa, C. (2015). Effects of genotype, environment, and their interactions on honey bee health in Europe. *Current Opinion in Insect Science, 10*, 177–184.

Meixner, M.D., Pinto, M.A., Bouga, M., Kryger, P., Ivanova, E., & Fuchs, S. (2013). Standard methods for characterising subspecies and ecotypes of *Apis mellifera*. In V. Dietemann, J. D. Ellis, & P. Neumann (Eds.), The COLOSS BEEBOOK, Volume I: standard methods for *Apis mellifera* research. *Journal of Apicultural Research, 52*(4), 1–28. doi:10.3896/IBRA.1.52.4.05

Miguel, I., Baylac, M., Iriondo, M., Manzano, C., & Garnery, L. (2011). Both geometric morphometric and microsatellite data consistently support the differentiation of the *Apis mellifera* M evolutionary branch. *Apidologie, 42*, 150–161.

Miguel, I., Iriondo, M., Garnery, L., Sheppard, W.S., & Estonba, A. (2007). Gene flow within the M evolutionary lineage of *Apis mellifera* : Role of the Pyrenees, isolation by distance and post-glacial re-colonization routes in the western Europe. *Apidologie, 38*, 141–155.

Mikheyev, A.S., Tin, M.M., Arora, J., & Seeley, T.D. (2015). Museum samples reveal rapid evolution by wild honey bees exposed to a novel parasite. *Nature Communications, 6*, 443.

Neumann, P., & Carreck, N.L. (2010). Honey bee colony losses. *Journal of Apicultural Research, 49*(1), 1–6. doi:10.3896/IBRA.1.49.1.01

Oleksa, A., Chybicki, I., Tofilski, A., & Burczyk, J. (2011). Nuclear and mitochondrial patterns of introgression into native dark bees (*Apis mellifera mellifera*) in Poland. *Journal of Apicultural Research, 50*(2), 116–129. doi:10.3896/IBRA.1.50.2.03

Oleksa, A., Wilde, J., Tofilski, A., & Chybicki, I.J. (2013). Partial reproductive isolation between European subspecies of honey bees. *Apidologie, 44*(5), 611–619.

Omega bio-tek. (2013). *E.N.Z.A Forensic DNA Kit* (Product manual).

Parejo, M., Wragg, D., Gauthier, L., Vignal, A., Neumann, P., & Neuditschko, M. (2016). Using whole-genome sequence information to foster conservation efforts for the European dark honey bee, *Apis mellifera mellifera*. *Frontiers in Ecology and Evolution, 4*(140), 1–15. doi:10.3389/fevo.2016.00140

Parker, R., Melathopoulos, A. P., White, R., Pernal, S.F., Guarna, M.M., & Foster, L.J. (2010). Ecological adaptation of diverse honey bee (*Apis mellifera*) populations. *PLoS ONE, 5*(6), e11096.

Pinto, A. M., Henriques, D., Chávez-Galarza, J., Kryger, P., Garnery, L., … Johnston, J. S. (2014). Genetic integrity of the Dark European honey bee (*Apis mellifera* mellifera) from protected populations: a genome-wide assessment using SNPs and mtDNA sequence data. *Journal of Apicultural Research, 53*(2), 269–278.

Pinto, M.A., Henriques, D., Chávez-Galarza, J., Kryger, P., Garnery, L., van der Zee, R., … Carreck, N. L. (2014). Genetic integrity of the dark European honey bee (*Apis mellifera mellifera*) from protected populations: A genome-wide assessment using SNPs and mtDNA sequence data. *Journal of Apicultural Research, 53*(2), 269–278. doi:10.3896/IBRA.1.53.2.08

Posada, D., & Crandall, K.A. (2001). Selecting the best-fit model of nucleotide substitution. *Systematic Biology, 50*(4), 580–601.

Pritchard, D. (2009). Is the dark bee really native to Britain and Ireland? *Bee Improvement and Conservation, 30*, 14–17.

Pritchard, J.K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959..

Qiagen. (2010). *Multiplex PCR handbook*.

Rortais, A., Arnold, G., Alburaki, M., Legout, H., & Garnery, L. (2011). Review of the DraI COI-COII test for the conservation of the black honey bee (*Apis mellifera mellifera*). *Conservation Genetics Resources, 3*(2), 383–391.

Ruottinen, L., Berg, P., Kantanen, J., Kristensen, T.N., & Praebel, A. 2014. *Status and conservation of the Nordic brown bee*.

Ruttner, F. (1988). *Biogeography and taxonomy of honey bees*. Berlin: Springer.

Soland-Reckeweg, G., Heckel, G., Neumann, P., Fluri, P., & Excoffier, L. (2009). Gene flow in admixed populations and implications for the conservation of the Western honey bee, *Apis mellifera*. *Journal of Insect Conservation, 13*(3), 317–328.

Soland-Reckeweg, G. (2006). *Genetic differentiation and hybridization in the honey bee (*Apis mellifera *L .) in Switzerland*. Bern: University of Bern.

Strange, J.P., Garnery, L., & Sheppard, W.S. (2008). Morphological and molecular characterization of the Landes honey bee (*Apis mellifera* L.) ecotype for genetic conservation. *Journal of Insect Conservation, 12*(5), 527–537.

Szabo, T.I., & Lefkovitch, L.P. (1988). Fourth generation of closed population honey bee breeding 2. Relationship between morphological and colony traits. *Apidologie, 19*(3), 259–274.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution, 30*(12), 2725–2729.

Tarpy, D. R. (2003). Genetic diversity within honey bee colonies prevents severe infections and promotes colony growth. *Proceedings. Biological sciences / The Royal Society, 270*(1510), 99–103.

Vähä, J.-P., & Primmer, C.R. (2006). Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology, 15*(1), 63–72.

Watson, J.K. (1981). *Bee-keeping in Ireland: A history*. Dublin: Glendale Press.

Williams, E. (2013). An examination of the distribution of native Irish honey bees (*Apis mellifera mellifera*). *The Four Seasons, 51*, 8–13.

SCIENTIFIC REPORTS

OPEN

High sample throughput genotyping for estimating C-lineage introgression in the dark honeybee: an accurate and cost-effective SNP-based tool

Received: 18 January 2018
Accepted: 17 May 2018
Published online: 04 June 2018

Dora Henriques[1,2], Keith A. Browne[3], Mark W. Barnett[4], Melanie Parejo[5], Per Kryger[6], Tom C. Freeman[4], Irene Muñoz[7], Lionel Garnery[8,9], Fiona Highet[10], J. Spencer Jonhston[11], Grace P. McCormack[3] & M. Alice Pinto[1]

The natural distribution of the honeybee (*Apis mellifera* L.) has been changed by humans in recent decades to such an extent that the formerly widest-spread European subspecies, *Apis mellifera mellifera*, is threatened by extinction through introgression from highly divergent commercial strains in large tracts of its range. Conservation efforts for *A. m. mellifera* are underway in multiple European countries requiring reliable and cost-efficient molecular tools to identify purebred colonies. Here, we developed four ancestry-informative SNP assays for high sample throughput genotyping using the iPLEX Mass Array system. Our customized assays were tested on DNA from individual and pooled, haploid and diploid honeybee samples extracted from different tissues using a diverse range of protocols. The assays had a high genotyping success rate and yielded accurate genotypes. Performance assessed against whole-genome data showed that individual assays behaved well, although the most accurate introgression estimates were obtained for the four assays combined (117 SNPs). The best compromise between accuracy and genotyping costs was achieved when combining two assays (62 SNPs). We provide a ready-to-use cost-effective tool for accurate molecular identification and estimation of introgression levels to more effectively monitor and manage *A. m. mellifera* conservatories.

Pollination by the honeybee (*Apis mellifera* L.) is a blended ecosystem service of managed and unmanaged (feral or wild) colonies that is under threat from human-mediated environmental changes including climate change, habitat loss, habitat fragmentation, pesticides, and introduced parasites and pathogens[1,2]. There is growing evidence that management of locally adapted genetic diversity in honeybee subspecies and ecotypes is key to the long-term sustainability of this service[3–5]. Accordingly, actions towards preserving the large stores of genetic diversity held by the 31 honeybee subspecies[6–9] are expected to counteract the trend of global colony losses.

Of the 31 subspecies that have been identified in the natural distributional range of *A. mellifera* in Africa, Middle East, Western Asia, and Europe[6,9,10] there are 10 European subspecies grouped into two evolutionary

[1]Mountain Research Centre (CIMO), Polytechnic Institute of Bragança, 5300-253, Bragança, Portugal. [2]Centre of Molecular and Environmental Biology (CBMA), University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal. [3]Department of Zoology, Ryan Institute, School of Natural Sciences, National University of Ireland Galway, Galway, Ireland. [4]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Edinburgh, Midlothian, EH25 9RG, Scotland, UK. [5]Agroscope, Swiss Bee Research Centre, 3003, Bern, Switzerland. [6]Aarhus University, Department of Agroecology, Slagelse, 4200, Denmark. [7]Área de Biología Animal, Dpto. de Zoología y Antropología Física, Universidad de Murcia, Campus de Espinardo, 30100, Murcia, Spain. [8]Laboratoire Evolution, Génomes et Spéciation, CNRS, Gif-sur-Yvette, France. [9]Saint Quentin en Yvelines, Université de Versailles, Versailles, France. [10]Science and Advice for Scottish Agriculture (SASA), Roddinglaw Road, Edinburgh, EH12 9FJ, Scotland, UK. [11]Department of Entomology, Texas A&M University, College Station, USA. Correspondence and requests for materials should be addressed to M.A.P. (email: apinto@ipb.pt)

lineages[10]: the Western and Northern European (lineage M) and the South-eastern European (lineage C). Lineage M includes only two subspecies: the Dark honeybee *Apis mellifera mellifera* and the Iberian honeybee *Apis mellifera iberiensis*. Yet, these two subspecies cover the largest territory in Europe with *A. m. iberiensis* occupying the Iberian Peninsula and *A. m. mellifera* ranging from France in the south to Scandinavia in the north, and from Ireland and the UK in the west to the Ural Mountains in the east[10]. Lineage C occurs in a smaller geographical area composed of the Apennine and Balkan peninsulas and includes the most widely kept honeybee subspecies: the Italian *Apis mellifera ligustica* and the Carniolan *Apis mellifera carnica*. In spite of its wide distribution, *A. m. mellifera* is the subspecies most under threat as it is considered extinct in many parts of Europe not only because of the human-mediated environmental changes but more insidiously through replacement by and introgression from non-indigenous subspecies, particularly *A. m. ligustica* and *A. m. carnica*[11–13].

It has been argued that, unlike with other domesticated stock organisms, management and selective breeding in honeybees increase genetic diversity through introgression[14]. However, this form of admixture reduces the frequency of locally adapted gene complexes, leading to an increased likelihood of reduced survival rates of colonies[15]. How to protect locally adapted gene complexes that are more suited to local environments is a growing problem, as the increased breeding and movement of C-lineage honeybees promotes sympatry and gene flow between *A. m. mellifera* and imported commercial breeds. Efforts to assist conservation of *A. m. mellifera* are gathering momentum in multiple European countries (www.sicamm.org) and with the knowledge that reduced adapted genetic diversity threatens both managed and unmanaged populations, the interests of commercial beekeeping and honeybee conservationists should be aligning, particularly in *A. m. mellifera* indigenous areas.

An important first step in protecting *A. m. mellifera* populations in official or unofficial conservatories is to give the stakeholders an accurate and cost-efficient tool to test for C-lineage introgression. Microsatellites have been extensively used to examine C-lineage introgression in *A. m. mellifera*[11,12,16]. Yet, the numerous advantages of SNPs over microsatellites promise to make them the tool of choice for population monitoring and conservation purposes. In addition to being more abundant and widespread in the genome[17], SNPs display lower genotyping error, have higher quality data, are more amenable to automated analysis and data interpretation, and can be easily transferred between laboratories[18]. Moreover, SNPs proved to be more powerful than microsatellites at estimating C-lineage introgression in *A. m. mellifera*[19]. These properties make SNPs a powerful tool for testing the breeding stock in *A. m. mellifera* conservatories and SNP data can be readily incorporated in shared genetic databases, facilitating implementation of a conservation strategy at the European scale.

Whilst SNP analysis on whole genome (WG) sequence data may be required in studies concerned with fine-scale relatedness, such deep sequencing is disproportionate when determining introgression levels for the discrimination of *A. m. mellifera* breeding stocks. Also, while costs have dropped dramatically, WG sequencing is still unaffordable for most stakeholders committed to the long-term sustainability and conservation of honeybees. Costs are accrued as WG analysis requires considerable computing storage and processing power and trained bioinformatics personnel. However, encouragingly, Muñoz, *et al.*[20] showed that reduced panels of highly informative SNPs can accurately identify honeybee stocks[20–23]. Genotyping using reduced SNP panels considerably decreases laboratory processing costs. Furthermore, analysis of the generated genotypes requires low computational power and conventional bioinformatics skills.

Muñoz, *et al.*[20] developed reduced SNP panels for genetic identification and introgression analysis in *A. m. mellifera*. The authors used a combination of metrics to rank by information content over 1183 SNPs that had been genotyped in *A. m. mellifera*, *A. m. ligustica* and *A. m. carnica* using the 1536-plex GoldenGate® Assay of Illumina[13]. The top-ranked SNPs were combined into five nested panels whose sizes (48, 96, 144, 192, 384 SNPs each[20]) fitted the plexes of the now discontinued GoldenGate® Assays formerly genotyped with the VeraCode® technology. Here, we built from the 144-SNP panel to propose four customized assays tailored for high sample throughput genotyping using the iPLEX MassARRAY system. By providing a ready-to-use molecular tool for accurately, rapidly, and cost-effectively genotyping large sample sizes of *A. m. mellifera*, we hope to bring affordable C-lineage introgression detection to stakeholders in the fight to safeguard remaining reservoirs of unique combinations of genes and adaptations in *A. m. mellifera* and to expand its reduced current distribution.

Results

Assay design, quality control and genotyping accuracy. Of the 144 highly-informative SNPs selected by Muñoz, *et al.*[20], the Assay Design software was able to multiplex 127 into four assays (identified by letter M), each containing a variable number of SNPs ranging from 38 in M1 to 24 in M4 (Supplementary Table S1). A total of 573 samples (Fig. 1) were genotyped for the four customized assays using the iPLEX MassARRAY system. Of the 573 samples, only seven displayed a SNP call failure rate >30% and these were excluded from further analysis (Supplementary Table S2). Of the 566 remaining samples, 551 displayed a low percentage (<10%) of missing data indicating a high genotyping success rate (96%).

The quality control and assessment of the genotyping accuracy of the 127 SNPs (Supplementary Table S3) led to identification of 10 problematic SNPs, of which seven were typed in <80% of the individuals, three were called heterozygous for >10% of the haploid individuals (Supplementary Tables S1 and S3), and three exhibited inconsistent calls among the three genotyping technologies in >5% of the individuals (Fig. 2). The latter SNPs were also identified as having high rates of missing data or heterozygosity (Supplementary Table S3). Once the 10 SNPs were removed from the datasets, the rates of missing data of the remaining 117 SNPs were low with 113 having <10% and four varying between 10.4% and 15.5% (Supplementary Table S1). The genotypes generated for the 117 high-quality SNPs in the MassARRAY platform were highly concordant with those of the Illumina's platforms (99.9% for the BeadArray and 99.6% for the HiSeq. 2500). Following the quality control step, 339 of the 573 genotyped samples had no missing data and the highest rate of missing data was 29% but only in two samples (Supplementary Table S2).

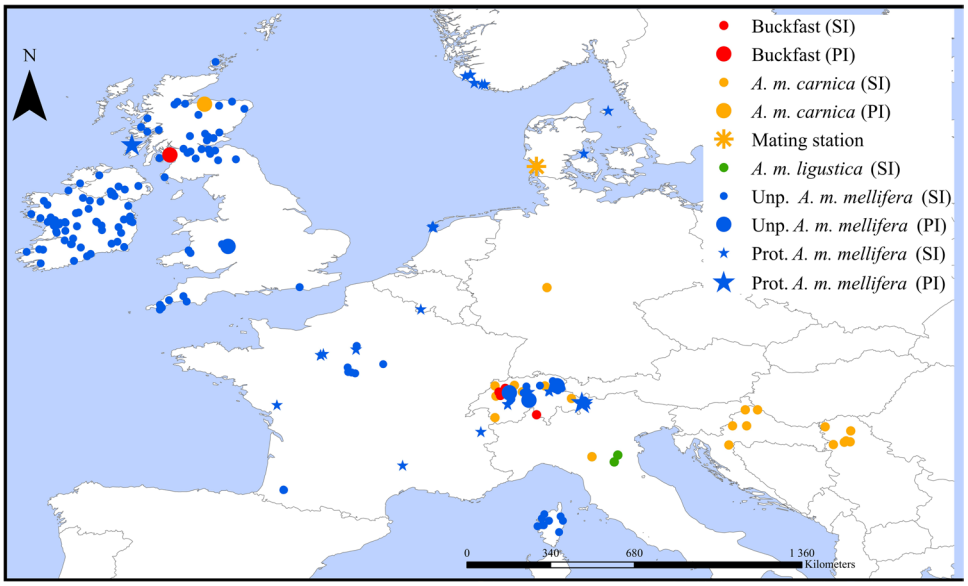


Figure 1. Location of the colonies sampled across the *A. m. mellifera* and C-lineage ranges. Samples of *A. m. mellifera* were collected in protected (Prot) and unprotected apiaries (Unp). The commercial breed Buckfast is also represented. Colonies were genotyped for the four SNP assays in the MassARRAY® MALDI-TOF platform from single individuals (SI) or pools of individuals (PI).

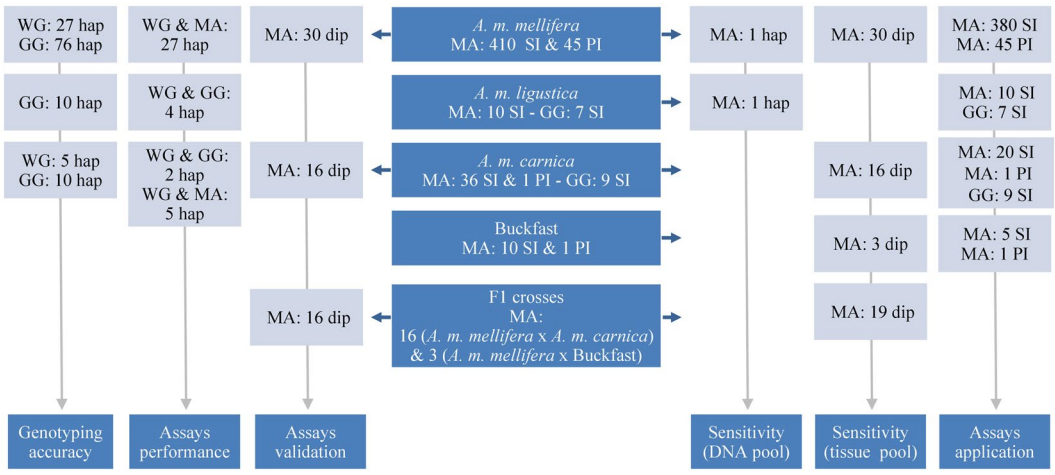


Figure 2. Datasets of quality-proved samples used in the SNP assays' testing and application. Samples were represented by a single individual (SI) or a pool of individuals (PI). The individuals were haploid drones (hap) or diploid workers (dip). Genotypes were generated from the four assays in the MassARRAY® MALDI-TOF platform (MA), from the GoldenGate® Assay in the Illumina's BeadArray platform (GG), and from whole genome (WG) sequences in the Illumina's HiSeq. 2500 platform.Vertical arrows connect the different individuals used in each test.

The final multiplexes contained M1 = 34, M2 = 32, M3 = 28, and M4 = 23 SNPs distributed across the 16 honeybee linkage groups, LGs (Fig. 3 and Supplementary Tables S1 and S4). LG 2 harboured the highest number of SNPs (13) while LG 3 had the lowest (2). The number of LGs covered by the assays varied between 12 (M4) and 14 (M1). Most SNPs (90 of 117) are located in non-coding regions, including intergenic (50 SNPs), intronic (30 SNPs), and UTRs (10 SNPs). Of the 27 coding SNPs, only two (1384-est6107 and 661-AMB-00398036) are non-synonymous (Supplementary Table S1).

Assessing performance of the SNP assays. The performance of the four assays was assessed by comparing their *Q*-values (inferred from single or combined assays) with those inferred from the genome-wide SNPs, which provides the best estimate of the admixture proportions (Supplementary Table S5). The four assays exhibited a good individual performance with a mean accuracy >94% and *Q*-values highly correlated ($0.980 \leq r \leq 0.983$) with those inferred from the WG dataset (Table 1). The largest plex assay M1 (34 SNPs) and

SNP Assay	# of SNPs	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
M1	34	0.983	0.929	0.046	0.211	26	95.42	0.061	2	1
M2	32	0.981	0.919	0.051	0.239	24	94.86	0.068	3	2
M3	28	0.982	0.926	0.047	0.314	23	95.27	0.066	3	0
M4	23	0.980	0.911	0.050	0.283	23	95.00	0.067	0	2
M1 + M3	62	0.993	0.956	0.029	0.172	31	97.09	0.042	2	0
M1 + M2 + M3	94	0.994	0.957	0.031	0.137	28	96.94	0.040	3	0
M1 + M2 + M3 + M4	117	0.996	0.964	0.022	0.114	32	97.84	0.033	2	0

Table 1. Statistics for the performance of the four SNP assays used singly or combined. Calculations were made via comparisons between *Q*-values inferred from the SNP assays and the genome-wide 2.399 million SNPs. (i) Pearson's correlation coefficient (*r*); (ii) similarity score obtained by CLUMPAK; (iii) mean and (iv) maximum absolute accuracy errors; (v) number of individuals (out of 38) with absolute accuracy error <0.05; (vi) mean accuracy estimated via percentage of absolute error; (vii) absolute precision error; (viii) number of purebred *A. m. mellifera* individuals misclassified as admixed; (ix) number of admixed individuals misclassified as purebred.

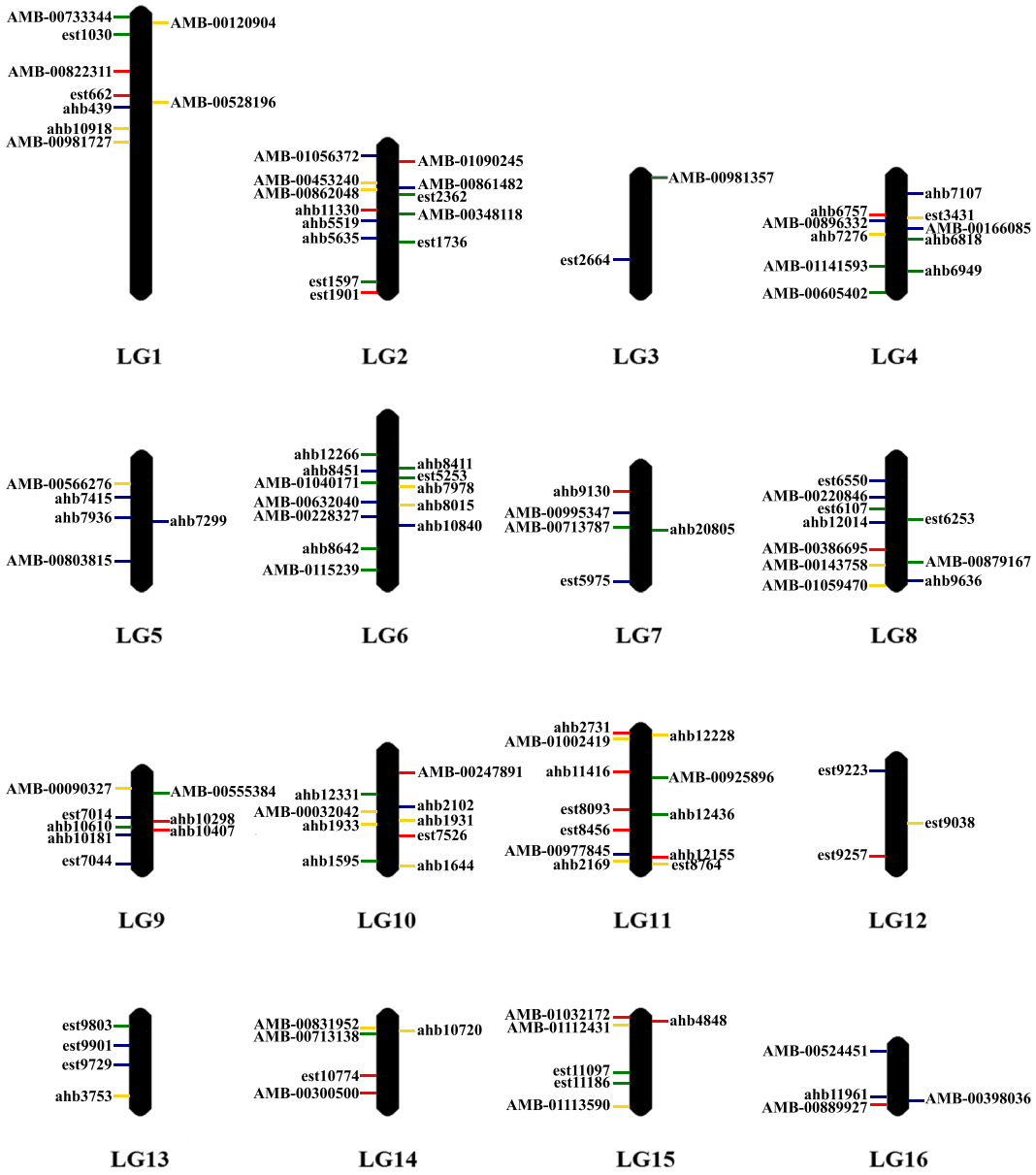


Figure 3. Genomic positions of the 117 quality-proved SNPs. The 117 SNPs were multiplexed in four assays, named M1 (blue), M2 (green), M3 (yellow), and M4 (red).

the smallest M4 (23 SNPs) showed the best and the worst behaviour, respectively, as indicated by most statistics (Table 1). The best performance was achieved when the four assays (117 SNPs) were used together ($r = 0.996$; mean accuracy = 97.84%; absolute precision error = 0.033), although the combination of M1 + M3 (62 SNPs) and M1 + M2 + M3 (94 SNPs) with the highest individual correlations produced equally interesting statistics with mean accuracies >96.9%, absolute precision error <0.04, and with over 28 individuals (out of 32) with absolute accuracy error <0.05. Performance was also assessed by counting purebred *A. m. mellifera* individuals misclassified as admixed (*Q*-values > 0.05) and *vice versa* (Table 1). Except for M4, single assays and their combinations repeatedly misclassified two or three (always identified amongst individuals M23, M24, M25, and M26; Supplementary Table S5) purebred as admixed from 11 *A. m. mellifera* individuals identified by genome-wide SNPs. The degree of *A. m. mellifera* misclassification was lower for the class "admixed identified as purebred" with M3, and its combination with one (M1), two (M1 + M2) or three assays (M1 + M2 + M4) correctly identifying all 16 admixed individuals ($0.05 < Q\text{-value} < 0.95$).

Validating the SNP assays. The assays were validated using an independent set of 62 individuals, including 30 *A. m. mellifera*, 16 *A. m. carnica*, and 16 F1 hybrids. On average, *Q*-values inferred from the genotypes called using the four individual (M1, M2, M3, M4) and three combined assays (M1 + M3, M1 + M2 + M3, M1 + M2 + M3 + M4) fit the thresholds defined for the two subspecies and hybrids ($P\text{-value} \geq 0.18$, Mann-Whitney test; Supplementary Table S6). Despite good overall performance of the individual assays, a few purebred *A. m. mellifera* and *A. m. carnica* were misclassified as admixed (estimated *Q*-values deviated from thresholds of <0.05 for *A. m. mellifera* and >0.95 for *A. m. carnica*) when the *Q*-values were inferred from called genotypes. However, when mixed combinations of the four assays were employed, the estimated *Q*-values matched the expectations with all *A. m. mellifera* and *A. m. carnica* correctly classified as purebred (Fig. 4a,b) and the F1 hybrids varying between 0.52 ± 0.04 (mean ± SD), for M1 + M3, and 0.56 ± 0.03, for the four assays combined, with the slight bias toward the *A. m. carnica* in the F1 (Fig. 4c) reflecting the known low level of C-derived introgression in the Læsø source population[13].

Assessing sensitivity of the MassARRAY system in pooled DNA. The sensitivity of the MassARRAY system in detecting *A. m. ligustica* was assessed in pools combining the DNA of two haploid individuals (one *A. m. mellifera* and one *A. m. ligustica*) at five dilution ratios. Of the 117 SNPs, only 103 were informative in this experiment (five were monomorphic, and nine were bi-allelic, but only one allele was called across dilutions). As expected, the sensitivity decreased as the dilution ratios increased, with only 29 unlinked SNPs being able to detect the *A. m. ligustica* alleles in every dilution and replicate (Supplementary Fig. S1, Supplementary Information). Yet, it was still possible to detect introgression with either the four assays (117 SNPs) or the two assays M1 + M3 (62 SNPs), even when the *A. m. ligustica* DNA was as diluted as 1:20 (Fig. 5 and Supplementary Table S7).

Assessing sensitivity of the MassARRAY system in pooled tissue. The sensitivity of the MassARRAY was further assessed in 22 tissue pools. Of the 2,574 called genotypes (117 SNP loci × 22 pools), 1,977 (77%) were accurate, as determined by comparing the calls for single workers with those of the pools. The most common sources of mismatch were "the most frequent allele" and "higher DNA concentration" (Table 2). The average rate of accurately called SNPs per pool was high (77%, 90 SNPs) and varied between 83% (97 SNPs), for the pools of two workers, and 50% (58 SNPs), for the pools of eight workers (Table 3).

The *Q*-values estimated for the 22 pools from the called genotypes were similar to those estimated from the expected genotypes using either the four assays, M1 + M3, or the 29 SNPs ($P\text{-value} \geq 0.35$, Mann-Whitney test). Furthermore, the MassARRAY platform was able to detect low frequency alleles, either of M-lineage (pools containing *A. m. mellifera*) or C-lineage ancestry (pools containing *A. m. carnica* or Buckfast), even when the tissue dilution was as low as 1:7 (Supplementary Table S8).

Applying the SNP assays. The four assays were applied to 478 colonies of various ancestries, represented by single (431 colonies) or pooled individuals (47 colonies), collected in 13 European countries (Fig. 1). The *Q*-values estimated for each *A. m. mellifera* colony (Fig. 6a), indicated that introgression varies throughout Europe, ranging on average from 0.0 in Norway to 0.447 ± 0.265 in Wales (Supplementary Table S9). The least introgressed *A. m. mellifera* colonies were from conservatories of Norway (0 ± 0.000), Scotland (0.006 ± 0.011) and Netherlands (0.046 ± 0.141) with over 80% of the individuals showing a $Q\text{-value} < 0.05$, although most individuals (91%) of the unprotected populations of Ireland were also very pure (0.021 ± 0.022). Populations of Denmark, France and Switzerland exhibited greater *Q*-values ($0.148 \leq Q\text{-value} \leq 0.280$) in both protected and unprotected populations with ≤11% of pure individuals. Admixture proportions estimated for *A. m. ligustica* and *A. m. carnica* sampled from native and introduced ranges showed that they are very pure ($0.972 \leq Q\text{-value} \leq 1.000$), excepting for some Swiss colonies (0.750 ± 0.296). The commercial breed Buckfast was mostly of C-derived ancestry (0.806 ± 0.055).

The genotype data were further examined by network analysis. The correlation network graph shown in Fig. 6b consisted of 5,522 edges and 418 nodes (samples). Samples with similar allele profiles clustered together. In total, three clusters were identified with cluster 1 containing 342 nodes (highest similarity to M-lineage), cluster 2 containing 58 nodes (highest similarity to C-lineage) and cluster 3 containing 18 nodes (highest rates of introgression). All samples from Norway, Ireland, Netherlands and Belgium were in cluster 1 whilst all samples from Italy, Croatia and Serbia were in cluster 2. Of 70 samples from Scotland, 61 samples were in cluster 1, 6 in cluster 2 and only 2 in cluster 3; a similar distribution was seen for samples from France and Switzerland. Samples from England, Denmark and Wales were also predominantly found in cluster 1.

The admixture patterns were also examined in pooled individuals representing an independent set of 47 colony samples from Switzerland and the UK (Supplementary Table S10). The average *Q*-values estimated for the

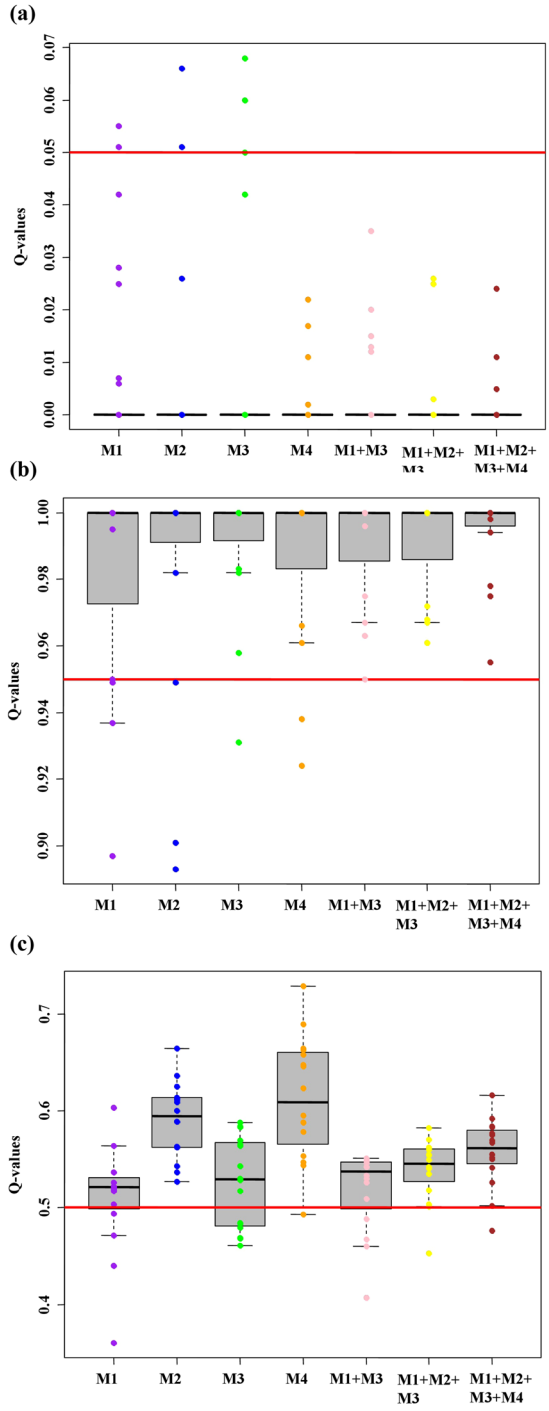


Figure 4. Validating the four SNP assays. Boxplots showing the variation of the *Q*-values inferred from the called genotypes for the four SNP assays. The boxes denote the first and third quartiles. The horizontal red lines mark the expected *Q*-values for purebred *A. m. mellifera and A. m. carnica* set at <0.05 and >0.95, respectively, and for the F1 hybrid samples set at 0.5. Boxplots for the (**a**) 30 *A. m. mellifera* samples, (**b**) 16 *A. m. carnica* samples, and (**c**) 16 F1 hybrid samples.

Swiss samples of *A. m. mellifera* varied between 0.145 ± 0.074 (protected) and 0.118 ± 0.042 (unprotected), which were lower than those inferred from a single individual (Supplementary Table S9). However, these estimates are not directly comparable as the pooled- and single-individual samples were from different apiaries. More comparable results were obtained for four colonies of variable ancestry from the UK that were simultaneously represented by a single worker and a pool of 16 workers. The *Q*-values inferred for each colony from the single worker and the pools were similar but always lower for the latter (Supplementary Table S10), a pattern that was also observed in the Swiss samples. This is an interesting finding that deserves to be fully investigated in a larger sample size.

SNP calling	# of genotypes
Sources of allele miscalling	
Different alleles	5
Higher DNA concentration	86
Higher DNA concentration & the most frequent allele	81
The most frequent allele	279
The least frequent allele	42
Missing data	104
Accurate calls	1,977
Total	2,574

Table 2. Information on SNP calling obtained from the 22 tissue pools.

Tissue pools (# of replicates)	Accurate SNPs	Miscalled SNPs				
		i	ii	iii	iv	v
1 Mel + 1 Hyb (3)	97.0	1.0	4.7	4.7	5.3	0.3
2 Mel + 1 Hyb (2)	81.0	0.0	3.5	5.0	21.0	0.5
3 Mel + 1 Hyb (2)	83.5	0.0	5.5	5.0	17.5	0.0
7 Mel + 1 Hyb (2)	58.0	0.0	4.5	5.0	47.0	0.5
1 Mel + 1 Car (3)	85.0	0.3	4.3	4.0	5.3	11.0
1 Car + 1 Hyb (3)	101.7	0.3	3.7	2.0	4.0	1.3
2 Car + 1 Hyb (2)	93.5	0.0	3.0	2.5	13.0	0.0
3 Car + 1 Hyb (2)	93.0	0.0	3.0	5.0	12.0	0.0
1 Buc + 1 Hyb (3)	102.7	0.0	3.0	1.3	4.7	0.7

Table 3. Mean number of SNP loci accurately called and miscalled for the different combination of tissue pools. The sources of miscalling were (i) different alleles, (ii) higher DNA concentration, (iii) higher DNA concentration and the most frequent allele, (iv) the most frequent allele, and (vi) the least frequent allele. Mel - *A. m. mellifera*; Hyb – F1 hybrid; Car – *A. m. carnica*; Buc – Buckfast.

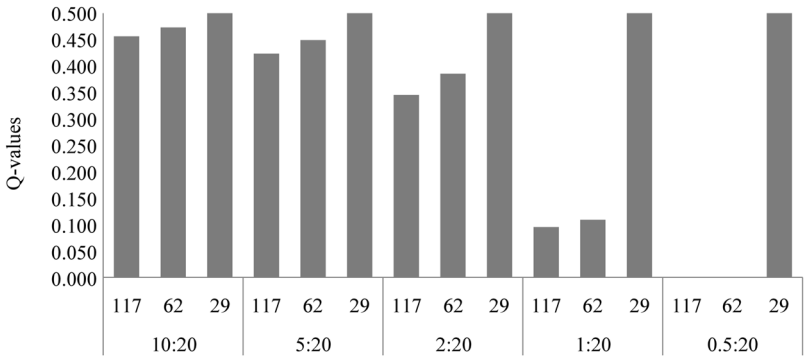


Figure 5. Average *Q*-values for different DNA pools. *Q*-values were inferred for DNA pools (representing dilution ratios of 10:20, 5:20, 2:20, 1:20, 0.5:20) by the four SNP assays (117 SNPs), the two best assays M1 + M3 (62 SNPs) and the 29 SNPs that were identified in all dilution ratios. The *Q*-values of the four SNPs assays and the M1 + M3 decreased as the dilution ratios increased. The *Q*-values of the 29 SNPs were always 0.50, as these were all heterozygous.

Discussion

The success of the numerous initiatives that are developing across Europe to protect and bring back the endangered dark honeybee rely on molecular tools capable of accurately detecting varying levels of C-derived introgression in a time- and cost-effective manner. In many conservation programs, the breeding stock has been routinely identified through wing morphometry and, more recently, through microsatellites[24]. However, inferring from data on Africanized honeybees[25], wing morphometry is likely unable to detect low levels of C-lineage introgression into *A. m. mellifera*, a limitation that is overcome by microsatellites[11,12]. While adoption of microsatellites represented a major step in conservation management of *A. m. mellifera*[12], it has been shown that a reduced number of high-graded SNPs[20], outperform the multiallelic marker in estimating introgression[19,26].

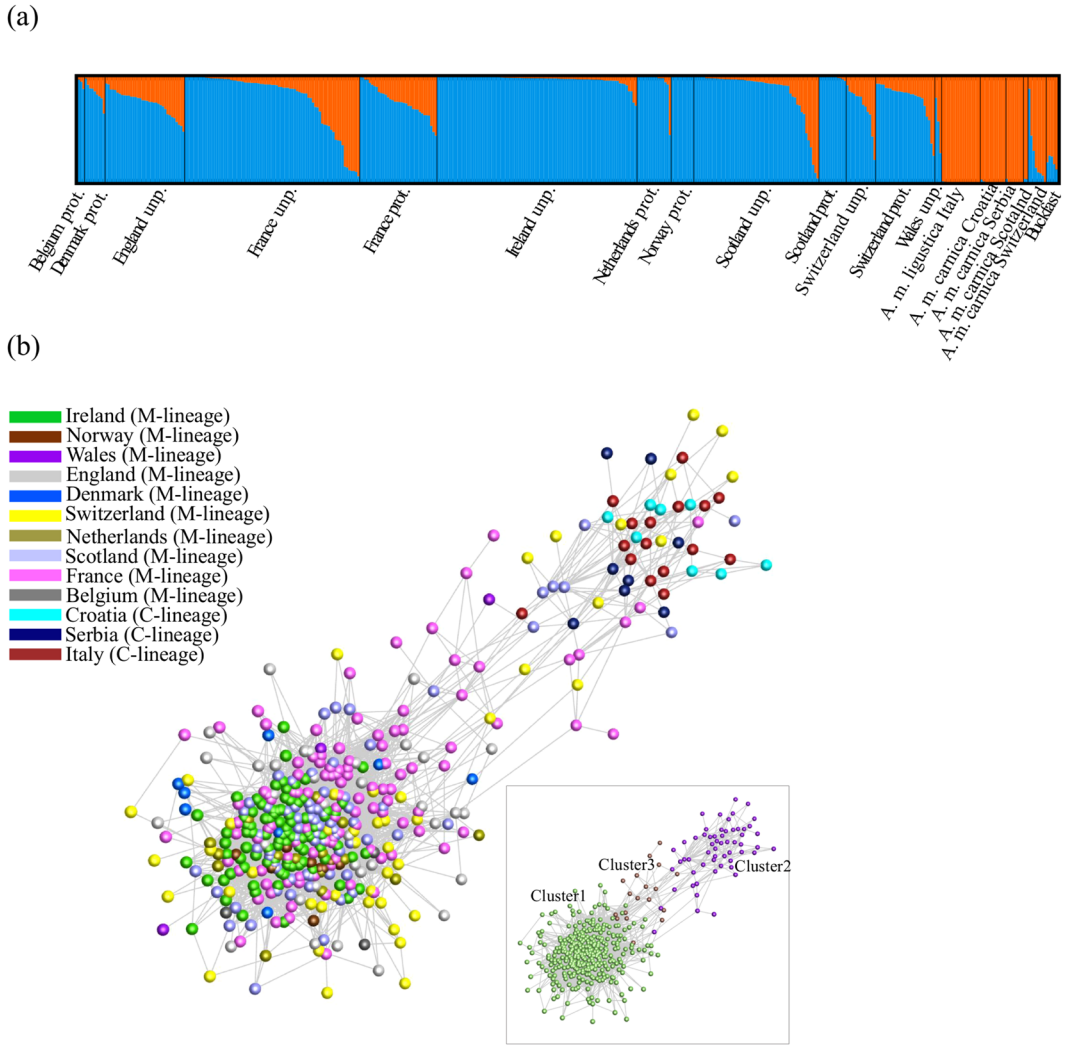


Figure 6. Structure reconstructed by ADMIXTURE and Graphia Professional software packages for honeybees of diverse ancestry collected across Europe. Most depicted samples (415) were genotyped in the MassARRAY platform using the four assays (117 SNPs). Nine samples of *A. m. carnica* and seven *A. m. ligustica*, previously genotyped for the 117 SNP loci using the GoldenGate Assay in the BeadArray platform, were added to the structure analysis for a better representation of C-lineage diversity. Each sample corresponds to a single colony. Samples collected in the *A. m. mellifera* range are from protected (prot) and unprotected (unp) apiaries. (**a**) ADMIXTURE plot showing the genome partitioning into two clusters (K = 2) for each individual, represented by a vertical bar. Blue represents the *A. m. mellifera* cluster and orange the C-lineage cluster. The black lines separate individuals from different countries and studied groups. (**b**) Correlation network where nodes (honeybee samples) are connected with edges when $r > 0.27$. A total of 418 samples out of 431 formed connections in the graph. Samples coloured according to country of origin with expected lineage indicated within parentheses. Inset shows correlation network clustered using the Markov Cluster (MCL) algorithm at an inflation value of 1.2.

Here, from the 144 top-ranked SNPs, selected by their power in discriminating C- from M-lineage honeybees[20], we designed, tested and validated four assays for genotyping with the iPLEX MassARRAY system. We provide the genomic information along with the PCR and iPLEX primers for 117 high-quality SNPs multiplexed in the four assays for immediate application in genetic surveys and conservation management of *A. m. mellifera*. In addition, we provide the dataset with the genotypes for haploid and diploid individuals of *A. m. mellifera*, *A. m. carnica* and *A. m. ligustica*, which can be used by others in introgression analysis as baseline reference populations with no need for inter-laboratory calibration[18]. As opposed to microsatellites, merging of SNP databases is straightforward as there are only two alleles per locus and different platforms will provide the same allele calls. If needed, curation will only involve SNP conversion from different platforms to be on the same DNA strand, which is much simpler than trying to harmonize different microsatellite allele sizes genotyped in different laboratories.

We show that C-lineage introgression can be accurately estimated from haploid, diploid, and combined haploid and diploid datasets (see Supplementary Information for details). These findings indicate that honeybee

conservation managers can choose the software of their preference and, more importantly, can simultaneously analyse workers and drones without biasing estimates of C-lineage introgression in *A. m. mellifera* colonies.

The Assay Design software was able to combine only 127 of the 144 high-graded SNPs[20] into four multiplexes. While the iPLEX protocol allows multiplexing up to 40 SNPs, only assay M1 (38 SNPs) approached the maximum plexing capacity. This is in part due to the relatively small size of the baseline SNP set from which the Assay Design had to work. However, the plex level of each assay can be expanded any time. By using the *Replex* option of the software, additional high-graded nuclear SNPs or even mitochondrial SNPs can be added to the customized four assays for detecting C-derived genes at both genetic compartments.

The iPLEX MassARRAY system revealed highly accurate and delivered high-quality calls for 117 of the 127 SNPs. Quality assessment was greatly facilitated by the honeybee haplodiploid system. Using the SNP calls of the drone subset, problematic SNPs were easily detected by locating genotypes erroneously typed heterozygous. Three such SNPs were consistently identified in numerous drones. While the mechanism responsible for the false allele is unclear, it is possible that gene homology is the source of miscalling at least in locus 1379-est5929. Using the 120-bp flanking region of this SNP locus, a NCBI query found a second hit with 98% similarity in the honeybee genome. The 117 SNPs were successfully genotyped in over 96% of the samples, indicating that the customized four assays and the iPLEX MassARRAY system work well in DNAs obtained from a variety of tissues with the virtually full spectrum of extraction methods routinely employed in honeybee research[27].

The four combined SNP assays were able to estimate introgression with a high degree of accuracy. However, performance decreased to some extent when SNP assays were used singly and the 23-plex M4 showed the worst behaviour for most statistics. This finding is consistent with studies on other organisms which have also detected drops in accuracy when the number of SNPs is <25[28,29]. Further assessment of the four assays (used singly or combined) at the individual level indicates that there is a greater chance of misclassifying purebred individuals as admixed than the reverse, *viz*. misclassifying admixed individuals as purebred. This result has practical implications in conservation management suggesting that it is more likely that *A. m. mellifera* genetic diversity is erroneously discarded from the breeding population than C-derived genes are maintained. At this point, simulation and empirical studies are needed to determine the best threshold criterion to separate purebreds from admixed individuals[30]. While the stringent *Q*-value threshold of <0.05 arbitrarily established here for defining purebred *A. m. mellifera* may assure a more efficient purging of C-derived alleles, it may also lead to erosion of *A. m. mellifera* diversity and loss of unique gene complexes. The problem is that low diversity is particularly detrimental for honeybees because it may decrease colony resistance to brood diseases[31] and increase genetic load at the sex locus[32]. Therefore, managers of *A. m. mellifera* conservatories need to make a trade-off between purging foreign alleles from the breeding population while minimizing the effects of reduced diversity.

Validation of the four SNP assays in an independent set of individuals, including F1 hybrids (obtained from controlled crosses purposely established for this study, as opposed to the simulated hybrids more commonly found in the literature), further confirms the resolution power of our customized SNP assays. Interestingly, the *Q*-values obtained for the F1 hybrids were in close proximity to the expected 0.50, although there was a bias towards C-derived genes as most *Q*-values were >0.50. When used singly, the SNP assays failed to correctly identify all purebred individuals and the *Q*-values were more dispersed around 0.50. However, when the assays were combined, the performance increased with all purebred individuals correctly classified and the *Q*-values showing a lower dispersion around 0.50. Interestingly, despite the lower number of SNPs contained in M1 + M3 (62 *vs* 117), this assay combination shows an overall performance similar to that of M1 + M2 + M3 + M4.

Sustainable conservation management requires tools capable of reliably identifying breeding colonies in a time- and cost-efficient manner. The SNP assays tested herein have a high resolution power for accurately estimating introgression, and the iPLEX MassARRAY system offers an interesting option for rapid and cost-effective genotyping. This system is very flexible and scalable allowing a variety of options for sample and assay throughput at a variable cost, depending on the chip format (24, 96, or 384) chosen. The 384 format, for example, allows genotyping 384 samples with a single assay at an approximate outsourced cost of 4.5€ per sample. Alternatively, this format could be used to genotype 192, 128, or 96 samples with two, three, or four assays, respectively. This option would incur in an increment of 4.5€ for any additional assay. Based on overall results, the best compromise between genotyping costs and assay accuracy is achieved when using M1 + M3.

Genotyping a single microsatellite multiplex in a 96-plate format costs approximately 2.5€ per sample. Introgression proportions using microsatellites has typically been estimated from over 11 loci, which requires genotyping a minimum of two multiplexes[11,24,26,33] thereby doubling the per-sample cost. However, this charge does not include PCR and microsatellite fragment analysis. Contrary to microsatellites, outsourced SNP genotyping with the iPLEX MassArray system only requires DNA (instead of PCR products) to generate a table of genotypes ready to analyse, avoiding the hurdle of fragment analysis.

Honeybee queens mate in flight with up to 20 drones[34]. This means that in areas where *A. m. mellifera* and commercial colonies are sympatric, matings may occur with drones of C-lineage ancestry originating colonies made up of subfamilies with diverse genetic backgrounds. Although population-level studies typically require genotyping a single worker per colony[16], colony-level introgression estimates may require genotyping several individuals to more effectively capture the colony structure. The problem is that genotyping several workers per colony is time consuming and costly. An economical way to circumvent this issue is to genotype pools instead of individuals[35], provided that the genotyping system of choice is sensitive enough to detect low-frequency alleles.

Here, we assessed whether our customized SNPs assays and the iPLEX MassARRAY system offer a reliable alternative for pool genotyping. Both DNA and tissue pooling experiments show that the genotyping system is very sensitive as it was able to detect low frequency alleles. Despite the small number of SNPs showing consistent amplification across experiments, introgression analysis indicates that as few as 62 SNPs (M1 + M3) were able to detect highly diluted C-derived alleles. These results suggest that this system has the potential to detect C-lineage

introgression in colonies with hybrid sub-families at low frequency, a scenario that might occur if drones of commercial colonies are able to accidentally enter congregation areas of conservatories.

Analysis of pool genotypes showed that miscalling was mainly due to the unequal contribution of each individual (different concentrations) and to the unbiased representation of allelic products that are present in a DNA pool, both common problems reported for DNA pools[35,36]. While pools constructed from equi-molar DNA concentrations would be the most correct approach to genotype a colony, pooling tissues is often the only option in conservation programs requiring screening of numerous colonies with a limited budget. Pooling tissue instead of DNA requires less time, effort and money during preparation in the laboratory and still enables detection of C-derived alleles even when most of the individuals in the pool are *A. m. mellifera*.

The introgression analysis on the samples collected throughout Europe and genotyped using the four SNP assays and the iPLEX MassARRAY system provides a rough picture of the genetic integrity of *A. m. mellifera*. This SNP survey adds to Pinto, *et al.*[13] by expanding the sampling in France, Switzerland, UK and by including *de novo* Wales and Ireland. Concordant with earlier microsatellite[11,12] and SNP[13,21] surveys, C-lineage introgression in *A. m. mellifera* is heterogeneous across Europe. Samples originating from conservatories were generally less introgressed than those from unprotected areas. Our previous and this SNP survey revealed that Scotland, Norway, Netherlands and now Ireland possess important pockets of pure *A. m. mellifera*. Ireland represents a particularly interesting case of *A. m. mellifera* diversity because, contrary to the other countries, the survey was performed in unprotected populations from a wide geographical area.

As this and previous studies[11–13,21] represent only partial, and in some cases biased, assessments on the status of the genetic integrity of *A. m. mellifera* across its distributional range, this novel tool now makes it possible to perform a comprehensive genetic survey in a time- and cost-efficient manner. We suggest that if the efficacy of this SNP tool is generally agreed among stakeholders the next step is for them to seek input from government agencies and/or research facilities and begin to describe the purity of their honeybee populations on as wide a geographic area as possible in order that conservation efforts correctly and efficiently target regions of greatest concern and greatest possible reward.

Methods

Assay design. Muñoz, *et al.*[20] identified 144 highly informative SNPs for estimating C-lineage introgression in *A. m. mellifera*. The flanking regions (60 bp of either side) of these SNPs were used to design multiplexed assays with the software Assay Design 4.0 (Agena BioScience™) for genotyping using the Agena BioScience iPLEX chemistry and the MassARRAY® MALDI-TOF platform (hereafter abbreviated to iPLEX MassARRAY). The software searched for optimal areas within the 120-bp flanking regions to design forward and reverse PCR primers while constructing the different multiplexes. The maximum multiplexing capacity (40 SNPs) allowed by the iPLEX chemistry was attempted whilst preventing hairpin and dimer formation. In addition to the PCR primers, the software designed the iPLEX extension primer placed immediately adjacent to each SNP. Of the 144 SNPs, the Assay Design was able to combine 127 SNPs distributed along four multiplexed assays (see Supplementary Table S1 for sequences of the flanking regions, sequences of PCR and iPLEX reaction primers, and composition of the four multiplexes). The putative functional role of the genes marked by each SNP was identified using SNPeff 4.3 tool build[37] and the NCBI *Apis mellifera* annotation genome version 102[38].

A wide array of analyses were carried out to validate the SNP assays and to assess (i) their accuracy when genotyped with the MassARRAY system, (ii) their performance when employed individually or combined, (iii) and their sensitivity when employed in pools of DNA and tissue. To that end, different combinations of samples, representing single and pooled, haploid and diploid individuals were used, as depicted in Fig. 2 and detailed in each section below.

Samples and DNA extraction. A total of 464 colonies (represented by a single haploid drone, a single diploid worker, multiple workers, or pools of drones or workers; Supplementary Table S2) were sampled across Europe (Fig. 1). The samples originated from colonies in the (i) *A. m. mellifera* (N = 462) native range in Western and Northern Europe (protected and unprotected areas), (ii) *A. m. ligustica* (N = 10), and *A. m. carnica* native ranges (N = 10) in South-eastern Europe, (iii) introduced range of *A. m. carnica* in Switzerland (N = 8), Germany (Kirchhain; N = 16) and Scotland (N = 3), (iv) commercial strain Buckfast from Switzerland, Scotland, and Denmark (N = 11), and (v) F1 hybrid crosses performed in isolated mating stations in Denmark (N = 19). Nine samples of *A. m. carnica* and seven *A. m. ligustica*, previously genotyped using the GoldenGate® Assay in the BeadArray platform of Illumina[13], were added to the dataset to have a better representation of C-lineage.

Genomic DNA was extracted from the head, antennae, thorax (entire or ~half), legs, or abdomen of adults or immatures (larvae or pupae) of a single individual, multiple individuals (extracted, then pooled), or a pool of individuals (mixed tissue, then extracted) per colony in 561 samples (Supplementary Table S2). The extraction methods included phenol-chloroform, CTAB, commercial kits (Qiagen EZ1 DNA tissue kit, Omega bio-tek EZNA kit), and magnetic beads using the KingFisher™ Flex Purification System. These represent the wide array of tissues and extraction methods commonly used in honeybee research[27]. The DNA samples were set at a concentration of 10–15 ng/µl and sent to *Instituto Gulbenkian de Ciência* (Portugal) for SNP genotyping.

SNP genotyping and quality control. A total of 573 samples (561 plus 12 DNA pools, Supplementary Table S2) were genotyped for the 127 SNP loci multiplexed in the four assays using the iPLEX chemistry and the MassARRAY® MALDI-TOF genotyping platform[39]. The genotypes generated for the 573 samples (Supplementary Table S11) were subjected to quality control filters to discard SNP loci and samples with poor or inconsistent amplification. SNPs and samples with missing data >20% (Supplementary Table S1) and >30% (Supplementary Table S2), respectively, were excluded from the dataset (Supplementary Table S3).

Assessing genotyping accuracy. The genotyping accuracy was assessed on the subset of single haploid drones of *A. m. mellifera* (N = 103), *A. m. ligustica* (N = 10) and *A. m. carnica* (N = 15), by (i) identifying the heterozygous SNP loci (N = 128; Fig. 2 and Supplementary Table S2) and (ii) comparing the SNP calls generated for a variable number of individuals by the iPLEX MassARRAY system with those obtained with the GoldenGate® Assay genotyped in the BeadArray platform of Illumina (N = 96 individuals[13]) and with the HiSeq. 2500 platform of Illumina (N = 32 individuals; see whole-genome sequencing details in Parejo, *et al.*[21] and Henriques, *et al.*[40]). The SNP loci that were called heterozygous by the MassARRAY system in >10% of the drones and showed inconsistent genotypes between at least two genotyping technologies in >5% of the drones were excluded from further analysis (Supplementary Tables S1 and S3).

Introgression estimation. Introgression proportions (*Q*-values) were estimated by ADMIXTURE[41] using datasets of varying ploidies (haploids, diploid, and their combination), which produced similar *Q*-values (see Supplementary Information for details). *Q*-values were estimated for K = 2 using 10,000 iterations in 20 independent runs. The convergence between iterations was monitored by comparing log-likelihood scores (LLS) using the default termination criterion set to stop when LLS increases by <0.0001 between iterations. CLUMPAK[42] was used to summarize and visualize the *Q*-plots.

Assessing performance of the SNP assays. The performance of the SNP assays in estimating C-lineage introgression in *A. m. mellifera* was assessed by comparing the *Q*-values inferred by them with those inferred from 2.399 million SNPs identified in WGs (see Parejo, *et al.*[21] and Henriques, *et al.*[40] for further details). A total of 38 drones (4 *A. m. ligustica*, 7 *A. m. carnica*, 11 purebred *A. m. mellifera*, and 16 admixed *A. m. mellifera*), for which there were WG sequence data available, was used in this comparison (Fig. 2). The 4 *A. m. ligustica* and 2 of the 7 *A. m. carnica* previously genotyped using the GoldenGate® Assay[13] were added to this step for a better representation of lineage C. The performance of the four assays (individually or combined) was assessed by (i) Pearson's correlation coefficient (*r*), (ii) similarity score obtained by CLUMPAK, (iii) absolute accuracy error calculated as the absolute difference between *Q*-values inferred from the SNP assays and the 2.399 million SNPs, (iv) mean accuracy calculated via percentage of absolute error, (v) absolute precision error calculated via standard deviation of the absolute differences, (vi) number of purebred individuals classified as admixed, and (vii) number of admixed individuals classified as purebred. Admixed individuals were defined by a threshold *Q*-value > 0.05. Any individual with *Q*-value between 0 and <0.05 or >0.95 and 1 was classified as purebred *A. m. mellifera* and C-lineage (*A. m.carnica* or *A. m. ligustica*), respectively.

Validating the SNP assays. The four assays were validated and tested using an independent subset of 62 workers, including 30 *A. m. mellifera* (Endelave, Denmark), 16 *A. m. carnica* (Kirchhain, Germany), and 16 F1 hybrids obtained from crosses between *A. m. mellifera* queens, from the conservatory in Læsø, and *A. m. carnica* drones from Mandø, Denmark (Fig. 2 and Supplementary Table S12). The crosses were performed in the isolated mating station of Mandø in 2016. *Q*-values were inferred from the four assays (individually or combined) by ADMIXTURE and then compared with the defined thresholds of >0.95 for *A. m. carnica*, <0.05 for *A. m. mellifera*, and ~0.5 for the F1 hybrids.

Assessing sensitivity of the MassARRAY system in pooled DNA. Pools of tissue or DNA are a cost-efficient option for estimating introgression in organisms with a polyandrous mating system like the honeybee. However, pooling can only be adopted if the genotyping system is able to consistently detect low-frequency alleles. The sensitivity of the MassARRAY system was assessed in a dilution experiment of varying ratios of DNAs of two haploid drones: one *A. m. ligustica* and one *A. m. mellifera* (Fig. 2). The two drones displayed the highest number of alternate alleles for the 127 highly-informative SNPs identified in a large dataset previously genotyped with the GoldenGate® Assay[13].

The experiment was performed by pooling the DNA of the two drones using volume ratios of 10:20, 5:20, 2:20, 1:20, and 0.5:20 *A. m. ligustica* to *A. m. mellifera* (Fig. 2). The number of replicates was three for 1:20 and 0.5:20 and two for the remaining ratios, as they were nested in the higher dilution factors. The pools were genotyped for the four assays using the iPLEX MassARRAY. The genotypes generated from the pooled DNAs were compared with those expected and the number of mismatches was recorded. The expected genotypes of the pools were inferred from the SNP calls for the single drones.

The sensitivity of the genotyping system in detecting C-lineage ancestry in the pooled samples was also assessed via introgression analysis. The *Q*-values were estimated by ADMIXTURE for each DNA pool using the expected and called genotypes for a variable number of SNPs (four assays and best assay combination, as defined by *r*).

Assessing sensitivity of the MassARRAY system in pooled tissue. The sensitivity of the MassARRAY system was further assessed in tissue pools (Supplementary Table S12). A total of 22 pools were constructed using varying ratios of workers (1:1, 1:2, 1:3, 1:7) of two different ancestries chosen among *A. m. mellifera* (N = 30), A. *m. carnica* (N = 16), Buckfast (N = 3), and F1 hybrids (*A. m. mellifera* queens x *A. m. carnica* drones; N = 19), as detailed in Fig. 2 and Supplementary Table S13. The DNA was extracted twice (individually and pooled) from the thorax, which had been cut in two identical portions. The DNA concentrations of individual and pooled extractions were measured using NanoDrop™ (Supplementary Table S12).

The sensitivity of the genotyping system was first assessed by comparing the SNP calls obtained for the single workers with those obtained for the pools of workers. Mismatches were counted and the error identified among the following sources: (i) pools displayed alleles uncalled in single workers and *vice versa*, (ii) SNP calls of the pools matched those of the worker with higher DNA concentration, (iii) SNP calls of the pools matched the most

frequent allele, and (iv) the least frequent allele. The sensitivity of the genotyping system in detecting C-lineage ancestry in the different pools was also assessed via introgression analysis. The *Q*-values were estimated for the 22 pools from the expected and called genotypes, for a variable number of SNPs (four assays and best assay combination), using ADMIXTURE. The expected genotypes were inferred from the calls obtained for the single workers.

Applying the SNP assays. The four assays were used to genotype in the MassARRAY platform 462 samples representing *A. m. mellifera* (N = 425), *A. m. ligustica* (N = 10), *A. m. carnica* (N = 21), and Buckfast (N = 6) from 8 13 European countries (Figs 1 and 2). Samples of *A. m. mellifera* originated from protected (N = 125) and unprotected (N = 300) areas. Of the 462 samples, 415 were represented by a single individual and 47 by pooled individuals (16 pooled workers from colonies of *A. m. mellifera*, *A. m. carnica* and Buckfast; 30 pooled drones from colonies of *A. m. mellifera*; Supplementary Table S2). Additionally, a subset of four colonies (two *A. m. mellifera*, one *A. m. carnica*, and one Buckfast) from Scotland and England was represented by both a pool of 16 workers and one individual worker. For a better C-lineage representation, nine samples of *A. m. carnica* and 7 of *A. m. ligustica* (each representing a single individual and colony), previously genotyped using the GoldenGate® Assay[13], were added to the dataset. *Q*-values were inferred from the genotypes of single and pooled samples using ADMIXTURE.

The genotype data were further examined by network analysis using the software Graphia Professional (Kajeka Ltd, Edinburgh, UK). For each sample, SNPs were scored 0 when same as reference (*A. m. carnica*), 1 for heterozygous and 2 for homozygous different to reference, i.e. representing the *A. m. mellifera* allele. Where data was missing, the SNP was scored 1.01. For ease of interpretation, the total combined score for each SNP in each sample was calculated and the SNPs reordered from the smallest score to the largest. The SNP data and associated sample metadata was loaded into Graphia and a Pearson correlation matrix was calculated comparing the profile of SNP scores for each sample. A network graph was then constructed by connecting the nodes (samples) with edges (where the correlation exceeded the threshold value $r > 0.27$). Utilising the overlay of metadata the graph was then explored and clustered using the Markov Cluster (MCL) algorithm[43] at an inflation value (which determines cluster granularity) of 1.2.

Data availability. *A. m. carnica* and *A. m. mellifera* whole-genome sequence data is deposited at the ENA (www.ebi.ac.uk/ena) under study accession number PRJEB16533.

References

1. Potts, S. G. *et al.* Global pollinator declines: trends, impacts and drivers. *Trends in Ecology & Evolution* **25**, 345–353, https://doi.org/10.1016/j.tree.2010.01.007 (2010).
2. van Engelsdorp, D. & Meixner, M. D. A historical review of managed honeybee populations in Europe and the United States and the factors that may affect them. *Journal of Invertebrate Pathology* **103**(Supplement), S80–S95, https://doi.org/10.1016/j.jip.2009.06.011 (2010).
3. Büchler, R. *et al.* The influence of genetic origin and its interaction with environmental effects on the survival of *Apis mellifera* L. colonies in Europe. *Journal of Apicultural Research* **53**, 205–214 (2014).
4. Meixner, M. D. *et al.* Occurrence of parasites and pathogens in honeybee colonies used in a European genotype-environment interactions experiment. *Journal of Apicultural Research* **53**, 215–229 (2014).
5. Meixner, M. D., Kryger, P. & Costa, C. Effects of genotype, environment, and their interactions on honeybee health in Europe. *Current Opinion in Insect Science* **10**, 177–184, https://doi.org/10.1016/j.cois.2015.05.010 (2015).
6. Chen, C. *et al.* Genomic analyses reveal demographic history and temperate adaptation of the newly discovered honeybee subspecies *Apis mellifera sinisxinyuan* n. ssp. *Molecular biology and evolution* **33**, 1337–1348 (2016).
7. Engel, M. S. The taxonomy of recent and fossil honeybees (Hymenoptera: Apidae; Apis). *Journal of Hymenoptera Research* **8** (1999).
8. Meixner, M. D., Leta, M. A., Koeniger, N. & Fuchs, S. The honeybees of Ethiopia represent a new subspecies of *Apis mellifera*- *Apis mellifera simensis* n. ssp. *Apidologie* **42**, 425–437, https://doi.org/10.1007/s13592-011-0007-y (2011).
9. Sheppard, W. S. & Meixner, M. D. *Apis mellifera pomonella*, a new honeybee subspecies from Central Asia. *Apidologie* **34**, 367–375, https://doi.org/10.1051/apido:2003037 (2003).
10. Ruttner, F. *Biogeography and Taxonomy of Honeybees* 165–257 (Springer-Verlag, 1988).
11. Jensen, A. B., Palmer, K. A., Boomsma, J. J. & Pedersen, B. V. Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honeybee, *Apis mellifera mellifera*, in northwest Europe. *Molecular Ecology* **14**, 93–106 (2005).
12. Soland-Reckeweg, G., Heckel, G., Neumann, P., Fluri, P. & Excoffier, L. Gene flow in admixed populations and implications for the conservation of the Western honeybee. *Apis mellifera. Journal of Insect Conservation* **13**, 317, https://doi.org/10.1007/s10841-008-9175-0 (2008).
13. Pinto, M. A. *et al.* Genetic integrity of the Dark European honeybee (*Apis mellifera mellifera*) from protected populations: a genome-wide assessment using SNPs and mtDNA sequence data. *Journal of Apicultural Research* **53**, 269–278, https://doi.org/10.3896/ibra.1.53.2.08 (2014).
14. Harpur, B. A., Minaei, S., Kent, C. F. & Zayed, A. Management increases genetic diversity of honeybees via admixture. *Molecular Ecology* **21**(4414–4421), 1365–4294X (2012).
15. De la Rúa, P. *et al.* Conserving genetic diversity in the honeybee: Comments on Harpur *et al.* (2012). *Molecular Ecology* **22**, 3208–3210, https://doi.org/10.1111/mec.12333 (2013).
16. Meixner, M. D. *et al.* Standard methods for characterising subspecies and ecotypes of *Apis mellifera*. *Journal of Apicultural Research* **52**, 1–28 (2013).
17. Weinstock, G. M. *et al.* Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* **443**, 931–949, https://doi.org/10.1038/nature05260 (2006).
18. Vignal, A., Milan, D., SanCristobal, M. & Eggen, A. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution* **34**, 275 (2002).
19. Muñoz, I. *et al.* SNPs selected by information content outperform randomly selected microsatellite loci for delineating genetic identification and introgression in the endangered dark European honeybee (*Apis mellifera mellifera*). *Molecular Ecology Resources* **17**, 783–795 (2017).
20. Muñoz, I. *et al.* Reduced SNP panels for genetic identification and introgression analysis in the dark honeybee (*Apis mellifera mellifera*). *Plos One* **10**, e0124365, https://doi.org/10.1371/journal.pone.0124365 (2015).
21. Parejo, M. *et al.* Using Whole-Genome Sequence information to foster conservation efforts for the european dark honeybee, *Apis mellifera mellifera*. *Frontiers in Ecology and Evolution* **4**, https://doi.org/10.3389/fevo.2016.00140 (2016).

22. Chapman, N. C. *et al*. A SNP test to identify Africanized honeybees via proportion of 'African' ancestry. *Molecular Ecology Resources* **15**, 1346–1355, https://doi.org/10.1111/1755-0998.12411 (2015).
23. Chapman, N. C. *et al*. An abbreviated SNP panel for ancestry assignment of honeybees (*Apis mellifera*). *Apidologie* **48**, 776–783 (2017).
24. Bouga, M. *et al*. A review of methods for discrimination of honeybee populations as applied to European beekeeping. *Journal of Apicultural Research* **50**, 51–84, https://doi.org/10.3896/ibra.1.50.1.06 (2011).
25. Guzmán-Novoa, E., Page, R. E. Jr. & Fondrk, M. K. Morphometric techniques do not detect intermediate and low levels of Africanization in honey bee (Hymenoptera: Apidae) colonies. *Annals of the Entomological Society of America* **87**, 507–515 (1994).
26. Parejo, M., Henriques, D., Pinto, M. A., S.-R., G. & Neuditschkoa, M. Empirical comparison of microsatellite and SNP markers to estimate introgression in *Apis mellifera mellifera*. *Journal of Apicultural Research* accepted (2018).
27. Evans, J. D. *et al*. Standard methods for molecular research in *Apis mellifera*. *Journal of Apicultural Research* **52**, 1–54 (2013).
28. Wilkinson, S. *et al*. Development of a genetic tool for product regulation in the diverse British pig breed market. *BMC Genomics* **13**, 580 (2012).
29. Storer, C. G. *et al*. Rank and order: evaluating the performance of SNPs for individual assignment in a non-model organism. *Plos One* **7**, e49018 (2012).
30. Vähä, J. P. & Primmer, C. R. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**, 63–72 (2006).
31. Seeley, T. D. & Tarpy, D. R. Queen promiscuity lowers disease within honeybee colonies. *Proceedings of the Royal Society of London B: Biological Sciences* **274**, 67–72 (2007).
32. Page, R. E. The evolution of multiple mating behavior by honeybee queens (*Apis mellifera* L.). *Genetics* **96**, 263–273 (1980).
33. Garnery, L. *et al*. Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*) II. Microsatellite loci. *Genetics Selection Evolution* **30**(S49), 1297–9686 (1998).
34. Estoup, A., Solignac, M. & Cornuet, J.-M. Precise assessment of the number of patrilines and of genetic relatedness in honeybee colonies. *Proceedings of the Royal Society of London B: Biological Sciences* **258**, 1–7 (1994).
35. Gautier, M. *et al*. Estimation of population allele frequencies from next-generation sequencing data: pool-versus individual-based genotyping. *Molecular Ecology* **22**, 3766–3779 (2013).
36. Sham, P., Bader, J. S., Craig, I., O'Donovan, M. & Owen, M. DNA Pooling: a tool for large-scale association studies. *Nature Reviews Genetics* **3**, 862–871 (2002).
37. Cingolani, P. *et al*. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strainw1118; iso-2; iso-3. *Fly* **6**, 80–92, https://doi.org/10.4161/fly.19695 (2012).
38. Pruitt, K. D. *et al*. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Research* **42**, D756–D763, https://doi.org/10.1093/nar/gkt1114 (2013).
39. Gabriel, S., Ziaugra, L. & Tabbaa, D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Current protocols in human genetics* 2.12. 11-12.12. 16 (2009).
40. Henriques, D. *et al*. Developing reduced SNP assays from whole-genome sequence data to estimate introgression in an organism with complex genetic patterns, the Iberian honeybee (*Apis mellifera iberiensis*). *Evolutionary Applications*, https://doi.org/10.1111/eva.12623 (2018).
41. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* **19**(1655–1664), 1088–9051 (2009).
42. Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* **15**(1179–1191), 1755–0998 (2015).
43. Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research* **30**, 1575–1584 (2002).

Acknowledgements

We are deeply indebted to João Costa (Instituto Gulbenkian Ciência, Oeiras, Portugal), for designing the multiplexes and SNP genotyping. José Rufino provided computational resources at the Polytechnic Institute of Bragança, Portugal. DH was supported by a PhD scholarship (SFRH/BD/84195/2012) from the Portuguese Science Foundation (FCT). KAB receives a PhD fellowship from the Irish Research Council. MP was supported by the Swiss Federal Office for Agriculture FOAG and the Fondation Sur-la-Croix, Basel. IM was supported by Saavedra Fajardo fellowship from the Fundación Séneca (20036/SF/16). MAP is a member of and receives support from the COST Action FA1307 (SUPER-B). Funding for genotyping of Irish honeybees was gratefully received from the Eva Crane Trust, the Native Irish Honeybee Society and the Department of Agriculture, Food and the Marine (16/GR/09). MB and TCF are funded by an Institute Strategic Grant from the Biotechnology and Biological Sciences Research Council (BBSRC) (BB/JO1446X/1). Financial support for this research was provided to MAP and LG by 2013–2014 BiodivERsA/FACCE-JPI joint call for research proposals, with the national funders FCT (Portugal), "Agence Nationale de la Recherche" (France), and "Ministério de Economia y Competividade" (Spain).

Author Contributions

M.A.P. and D.H. conceived the ideas and designed methodology with input from P.K. in the crosses and tissue pools; D.H. performed most of the analyses with assistance of K.A.B., M.P. and I.M.; M.W.B. and T.C.F. performed the network analysis. P.K. made the F1 crosses and constructed the tissue pools. M.A.P., K.A.B., M.W.B., M.P., P.K., T.C.F., L.G., F.H. and G.P.M. contributed with honeybee samples. J.S.J. provided the flanking sequences of the 144 SNPs. M.A.P. and D.H. wrote the manuscript with input from K.A.B. All the authors critically reviewed the manuscript for important intellectual content.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-26932-1.

Competing Interests: The authors declare no competing interests.

Journal of Apicultural Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjar20

Investigation of free-living honey bee colonies in Ireland

Keith A. Browne , Jack Hassett , Michael Geary , Elizabeth Moore , Dora Henriques , Gabriele Soland-Reckeweg , Roberto Ferrari , Eoin Mac Loughlin , Elizabeth O'Brien , Saoirse O'Driscoll , Philip Young , M. Alice Pinto & Grace P McCormack

To cite this article: Keith A. Browne , Jack Hassett , Michael Geary , Elizabeth Moore , Dora Henriques , Gabriele Soland-Reckeweg , Roberto Ferrari , Eoin Mac Loughlin , Elizabeth O'Brien , Saoirse O'Driscoll , Philip Young , M. Alice Pinto & Grace P McCormack (2020): Investigation of free-living honey bee colonies in Ireland, Journal of Apicultural Research, DOI: 10.1080/00218839.2020.1837530

To link to this article: https://doi.org/10.1080/00218839.2020.1837530

Published online: 03 Nov 2020.

Journal of Apicultural Research, 2020
https://doi.org/10.1080/00218839.2020.1837530

ORIGINAL RESEARCH ARTICLE

Investigation of free-living honey bee colonies in Ireland

Keith A. Browne[a], Jack Hassett[b], Michael Geary[b], Elizabeth Moore[b], Dora Henriques[c], Gabriele Soland-Reckeweg[d], Roberto Ferrari[a], Eoin Mac Loughlin[a], Elizabeth O'Brien[a], Saoirse O'Driscoll[a], Philip Young[a], M. Alice Pinto[c] and Grace P McCormack[a]*

[a]*Department of Zoology, Ryan Institute, School of Natural Sciences, National University of Ireland Galway, Galway, Ireland;* [b]*Department of Applied Science, Limerick Institute of Technology, Limerick, Ireland;* [c]*Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, Bragança, Portugal;* [d]*Apigenix, Twann, Switzerland*

(Received 23 April 2020; accepted 25 September 2020)

Apis mellifera mellifera (Linnaeus), the Western European honey bee, is considered extinct in the wild over most of its range due largely to hybridisation and replacement by other subspecies, parasitism by *Varroa destructor,* habitat loss, and effects from agricultural pesticides. The purity of the subspecies within the managed cohort is also at risk over much of its range. Here, we investigated if honey bee colonies inhabited locations outside of the apiaries. In those we located, we explored how long the colony persisted and we investigated the genotypes of the bees using multiple markers. We show here that unmanaged free-living honey bee colonies are present and widespread in Ireland, inhabiting a mixture of nesting habitats with some colonies persisting naturally and unaided over multiple years. Molecular data including mitochondrial, microsatellite, and SNPs evidence indicate that the free-living population sampled is largely comprised of pure *A. m. mellifera*. Finally, we discuss the implications of conserving free-living *A. m. mellifera* in Ireland and its possible role in improving the fitness of the managed population both in Ireland and the rest of its European range.

Keywords: *Apis mellifera*; subspecies; *Varroa destructor*; wild bees; feral bees; conservation; molecular data; survival

Introduction

Some subspecies of *Apis mellifera* (Order Hymenoptera, Family Apidae), e.g., *A. m. mellifera*, are currently in a state of near extinction across much of their range due to multiple factors including varroosis caused by the mite *Varroa destructor* (Order Mesostigmata, Family Varroidae) (Anderson & Trueman, 2000), changes in land use and the proliferation of pesticide use. An additional factor creating risk for the continued survival of native subspecies is the replacement of native with imported honey bee strains by beekeepers. The presence of non-native breeding stock and hybrid strains such as *A. m. carnica* and *A. m. ligustica* (C-lineage) and cross-lineage commercial hybrids such as 'Buckfast' has resulted in large-scale introgression between these and native bees reducing the population of pure *A. m. mellifera* and altering its genetic integrity, leading to the strong possibility that genes for locally adapted traits may have been removed from the population (De la Rua et al., 2009; Ellis et al., 2018; Jensen et al., 2005; Parejo et al., 2018; Pinto et al., 2014; Randi, 2008; Soland-Reckeweg, 2006). In the midst of efforts to address such issues mentioned above in the managed honey bee cohort there has been insufficient investigation into the status of wild honey bees leaving considerable uncertainty about their current state (e.g., abundance, distribution, longevity) and conservation need (Moritz et al., 2005; Nieto et al., 2014).

Wild *A. m. mellifera* colonies are thought to be extinct or near extinct in all but a handful of conservation areas and nature reserves (Kohl & Rutschmann, 2018; Moritz et al., 2007). In Ireland, few wild colonies have been seen by experienced beekeepers in recent times (e.g., Micheal MacGiolla Coda, Personal Communication). Though members of the public sometimes reported the presence of bees in chimneys to the authors and to beekeeping associations they were assumed to be escapees from local apiaries and likely hybrids. However, Jaffe et al. (2010) indicated that there were more colonies present in Ireland than could be accounted for by managed colonies only using a molecular approach. For all other countries sampled in that study (apart from one location in Italy), the numbers of colonies estimated from the genotyping approach matched with the number of managed colonies documented, indicating a loss of wild colonies at these locations (Jaffe et al., 2010). We thus sought to investigate if honey bees live under wild conditions in Ireland and if so to determine their status in terms of habitat choice, survival and genetic diversity. Due the difficulty with knowing if bees are truly wild (descended from colonies that never inhabited a beehive) or feral (having swarmed from a nearby apiary) we will use the term 'free-living' to define those bees that have been collected from any

*Corresponding author. Email: Grace.mccormack@nuigalway.ie

habitat excluding a beehive (it includes wild and feral colonies). We considered that this study was essential to inform stakeholders in honey bee conservation both in Ireland and worldwide given that the conservation of native and wild honey bee populations, and local ecotypes, are increasingly considered of high importance (e.g., Requier et al., 2019) particularly those that may be Varroa tolerant/resistant and that pure *A. m. mellifera* has been reported to be common in Ireland (Hassett et al., 2018; Henriques et al., 2018).

Materials and methods

Colony sampling and monitoring

In November 2015 and August 2016 a nationwide appeal for information on the location of free-living honey bee colonies was made via The Irish Times, a national newspaper. Reports were gathered by telephone, email, and social media contact, and sightings determined to be honey bee swarm capture or other bee species were excluded. Samples of honey bees were collected from entrances of all 76 colonies both validated as being free-living honey bee colonies at the time of sampling and accessible for sampling (Figure 1) using a combination of long-handled butterfly nets, clear glass jars, and a proprietary "Bug buster" suction tube. Samples were cooled immediately upon capture before storage at −20 °C until DNA extraction. Colony survival was observed by determining the presence or absence of activity in Spring and late Autumn of each year, and by identifying the cause of death where possible/applicable.

DNA extraction, PCR and sequencing

DNA was extracted from the two hind legs of worker bees using the E.Z.N.A. Forensics DNA extraction kit (Omega Bio-Tek). Mitochondrial DNA consisting of the highly polymorphic intergenic region between the 3′ end of the $tRNA^{leu}$ gene and the 5′ end of the COII subunit gene was amplified using E2 (5′-GGCAGAATAAGTGCATTG-3′) and H2 (5′-CAATATCATTGATGACC-3′) primers (Garnery et al., 1998) with Illustra PuReTaq Ready-To-Go PCR Beads (GE Healthcare). Polymerase chain reactions (PCR) included an initial denaturation of 5 mins at 95 °C, followed by 35 cycles of 94 °C for 45 secs, 45 °C for 43 secs and 62 °C for 2 mins with a final extension of 20 min at 65 °C (Garnery et al., 1993). PCR products were purified using a GeneJet PCR Purification kit (Thermo scientific) and sequenced by LGC Genomics, Germany. Resulting sequences were manually assessed against their chromatographs in MEGA7 (Kumar et al., 2016) before being imported into a multiple alignment.

Mitochondrial DNA sequence analysis

A total of 99 sequences of satisfactory quality were generated from 49 of the 76 free-living colonies (Table 1). Previous work (Hassett et al., 2018) indicated that all sequences from the free-living bees were of the M lineage. Here we explore further the relationships of the free-living bees using networks as implemented in TCS 1.21 (Clement et al., 2000). Firstly a network was produced from an alignment of the Irish free-living honey bee sequences with related sequences from Genbank and 156 sequences from the Irish managed honey bee cohort (Hassett et al., 2018; Pinto et al., 2014). The extra Q elements present in some haplotypes have a large influence on determining clusters given their length therefore an additional alignment was analysed that included only the informative sites in each Q element (showing some variation between bees). All unique sequences from the free-living honey bee colonies have been deposited into GenBank (accession numbers MT823282-MT823299).

Microsatellites analysis

Genotyping using a twelve microsatellite panel (A273, A43, Ac306, Ap33, B24, Ap226, A76-2p, A007, Ap001, A28, Ap289 and A29) was carried out by Ecogenics, Switzerland. Satisfactory data (ie. low amounts of null alleles) was returned from 59 of the free-living colonies as indicated in Table 1. This includes data from Hassett *et al* (2018) plus 24 additional free-living colonies (total N = 94). Equivalent data (the same loci, and carried out by the same laboratory using the same standards) were included from reference populations of *A. m. ligustica* (Soland-Reckeweg et al., 2009) from Italy (n = 55); *A. m. carnica* from Austria (n = 182) and Slovenia (n = 21), *A. m. mellifera* from Sweden (n = 10), France (n = 24), Norway (n = 18), Switzerland (n = 22) and managed bees from Ireland (n = 171) (Hassett et al., 2018). Bayesian analysis and visualisation of population assignment between C and M lineages was conducted in STRUCTURE V2.3.4 (Pritchard et al., 2000) using the admixture and correlated allele frequency models with the unsupervised option. A total of 750,000 Markov chain Monte Carlo (MCMC) iterations after an initial burn-in of 250,000 were performed for 20 iterations of each of K = 1 to 6. The optimal value of K (Evanno et al., 2005) was calculated using the CLUMPAK (Kopelman et al., 2015) online calculator. Nine populations were designated prior to analysis; seven based on the reference populations as above, and two for the Irish population divided between managed and free-living cohorts (1: Italy ligustica, 2: Austrian carnica, 3: Slovenian carnica 4: Swedish mellifera, 5: French mellifera, 6: Norwegian mellifera, 7: Swiss mellifera 8: Irish managed and 9: Irish free-living). The Q-value threshold used for full assignment to a particular population was $\geq$ 0.900 (Vaha & Primmer, 2006). The populations were

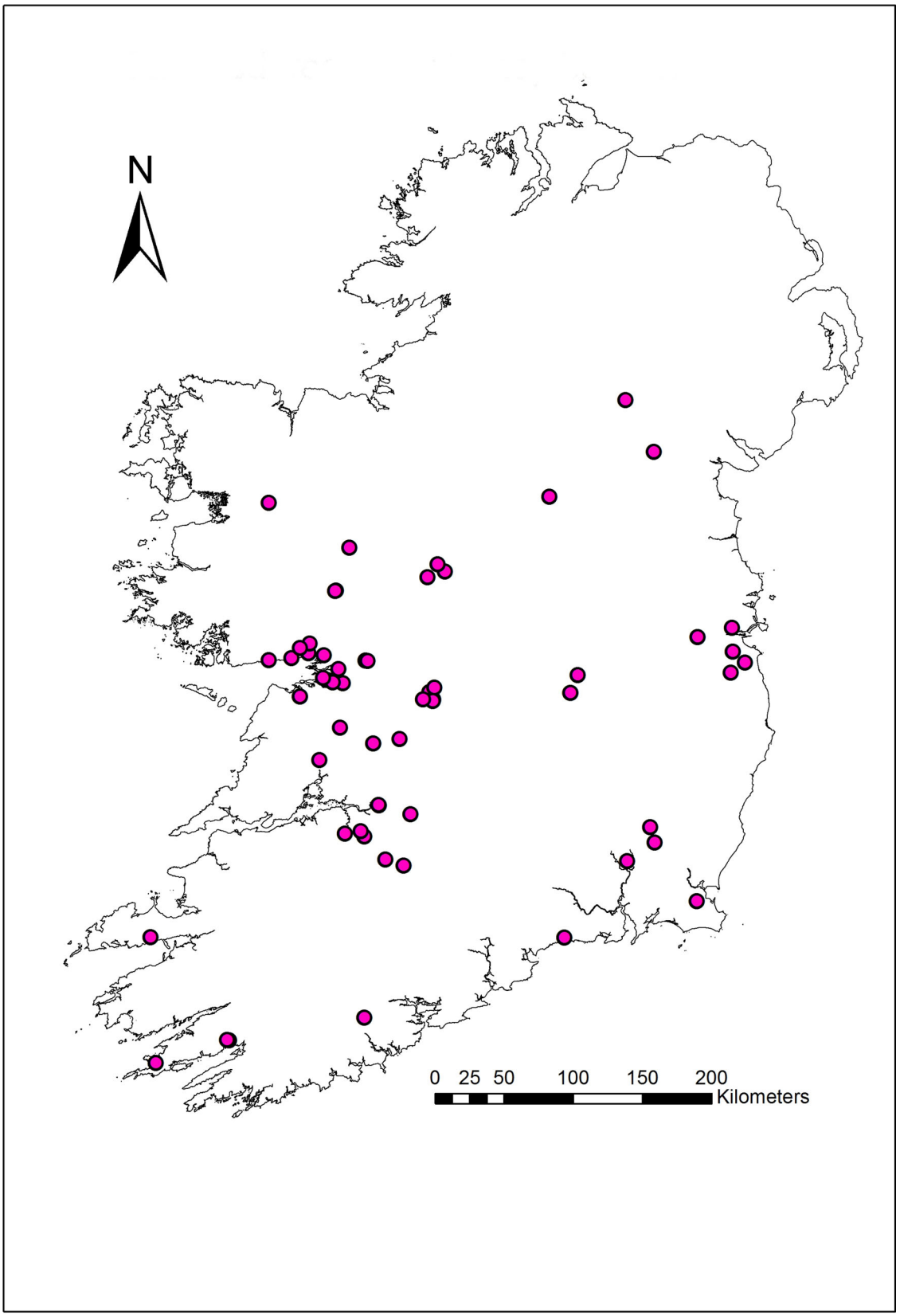


Figure 1. Locations of the 76 free-living colonies sampled in Ireland between 2015 and 2018 from the 182 reported. Some locations contained multiple colonies.

Table 1. Details of all free-living honey bee colonies sampled including information on their indicated lineage (M versus C) via mitochondrial, microsatellite and SNP data.

	mtDNA		Msats		SNPs	
Colony ID	**♯bees**	**Lineage**	**♯ bees**	**Prop M**	**♯ bees**	**Prop M**
F1T	nd	nd	nd	Nd	1	0.978
F2L	1	M	2	0.997-8	1	0.989
F3CE	nd	nd	nd	Nd	9	0.999
F4L	2	M	3	0.996-8	1	1
F5LA	nd	nd	3	0.998	1	0.991
F5LB	1	M	nd	nd	nd	nd
F6L	1	M	2	0.996	nd	nd
F7L	1	M	1	0.997	nd	nd
F8KEA	nd	nd	2	0.992-5	1	0.991
F8KEB	1	M	1	0.998	1	0.980
F19DA	1	M	1	0.998	1	1
F19DB	1	M	1	0.981	1	0.976
F10G	1	M	1	0.997	1	1
F11R	1	M	nd	Nd	nd	Nd
F12R	1	M	2	0.998	1	0.980
F13R	1	M	2	0.989-0.993	1	1
F14G	2	M	1	0.998	9	1
F15G	2	M	1	0.989	nd	Nd
F16G	1	M	1	0.997	9	0.983
F17G	nd	nd	nd	nd	nd	Nd
F18G	1	M	nd	nd	1	0.968
F19D	1	M	nd	nd	1	1
F20CE	1	M	2	0.998	9	0.984
F21C	**5**	**M**	**4**	**0.607-0.998**	**1**	**0.728**
F22	1	M	1	0.996	nd	Nd
F23	2	nd	nd	nd	nd	Nd
F24G	2	M	nd	nd	nd	Nd
F25G	1	M	nd	nd	nd	Nd
F26g	nd	nd	nd	nd	nd	Nd
F27G	2	M	2	0.9980	9	1
F28G	1		2	0.9980	1	1
F29G	7	M	5	0.933-997	8	1
F30G	1	M	1	0.9980	9	1
F31G	2	M	1	0.9900	nd	Nd
F32WX	nd	nd	nd	nd	1	0.974
F33KY	nd	nd	1	0.9980	1	1
F34W	2	M	1	0.9960	nd	Nd
F35CW	1	M	nd	nd	nd	Nd
F36C	2	M	1	0.9970	nd	Nd
F37C	nd	nd	1	0.991	1	0.9903
F38C	2	M	1	0.997	nd	Nd
F39G	nd	nd	nd	nd	nd	Nd
F40G	nd	nd	nd	nd	nd	Nd
F41G	3	M	2	0.993-4	9	1
F42G	nd	nd	1	0.997	nd	Nd
F43G	1	M	1	0.998	nd	Nd
F44G	7	M	5	**0.893-0.997**	nd	Nd
F45G	2		2	0.998	9	0.988
F46G	nd	nd	nd	nd	nd	Nd
F47G	3	M	2	0.989-0.998	9	1
F48G	8	M	4	0.997-0.998	9	1
F49R	1	M	1	0.994	nd	Nd
F50R	2	M	1	0.989	nd	Nd
F51MO	1	M	1	0.996	1	0.994
F52D	1	M	1	0.986	1	0.967
F53LO			1	0.997	1	0.985
F54MN	2	M	2	0.986-0.997	1	0.943
F55MN	2	M	2	**0.821, 0.988**	1	0.967
F56LS	2	M	1	0.994	1	0.968
F57OY	2	M	1	0.997-8	1	0.991
F58C	nd	nd	1	0.9960	nd	Nd
F59L	nd	nd	1	0.9980	nd	Nd
F60L	nd	nd	1	0.9920	nd	Nd

(*Continued*)

Table 1. (*Continued*).

	mtDNA		Msats		SNPs	
F61L	5	M	4	0.995-8	nd	Nd
F62L	nd	nd	1	0.9940	nd	Nd
F63WX	nd	nd	1	0.9980	nd	Nd
F64CE	nd	nd	1	0.9740	nd	Nd
F65OY	nd	nd	1	0.9970	nd	Nd
F66G	nd	nd	1	0.9160	nd	Nd
F67G	1	M	1	0.9730	nd	Nd
F68G	1	M	1	0.9980	nd	Nd
F69WW	nd	nd	1	0.9970	nd	Nd
F70CE	nd	nd	1	0.9850	nd	Nd
F71G	1	M	nd	nd	nd	Nd
F72G	1	M	2	0.9970	nd	Nd
F73G	1	M	1	0.9970	nd	Nd
76	**99**	49	**94**		**123**	36

Note: ♯= number of, Msat = microsatellite, SNP = Single Nucleotide Polymorphism, Prob M = probability of individual bee being part of M lineage. Colonies highlighted in grey are those that have results from all three types of data. The three colonies in bold are those where one or more data-type indicate introgression of C into M lineage.

arranged into four groups; European C-lineage (populations. 1, 2 and 3), European mellifera (populations. 4, 5, 6 and 7), Irish managed (population 8) and Irish free-living (population 9). Analysis was carried out both with all sampled bees and also with only one bee per colony to remove bias that may be caused by the presence of related bees. Using the groups above and the reduced dataset of one representative bee per colony for Irish colonies (Irish managed cohort N = 64, free-living bee cohort, N = 56) an analysis of molecular variance (AMOVA) was conducted using Arlequin V3.5.2.2 (Excoffier & Lischer, 2010).

Single nucleotide polymorphisms (SNPs) analysis

As a further approach to investigate the subspecies purity of the free living bees, DNA from 127 free-living bees, representing 39 colonies, were diluted to 10–15 ng/μl and sent to Instituto Gulbenkian de Ciência (Portugal) for genotyping using the Agena BioScience iPLEX chemistry and the MassARRAY® MALDI-TOF platform (Gabriel et al., 2009) using a highly informative 127 SNP assay, designed for reliable introgression estimation of C-into M-lineage (Henriques et al., 2018). After quality control to identify SNPs with missing data, data from 36 colonies were kept in the final analysis (Table 1). While for most colonies one bee was tested for introgression, for 11 colonies nine individual bees were tested and for one colony eight bees were tested (Table 1). Membership proportions (Q-values) were estimated using ADMIXTURE V1.23 (Alexander et al., 2009) for K = 2 with 20 independent runs of 10,000 iterations. The convergence between iterations was examined by comparing log-likelihood scores (LLS) using the default termination criteria set to stop when LLS increases by <0.0001 between iterations. A total of 36 M-lineage and 36 C-lineage individuals were used as reference populations (Henriques et al., 2018). CLUMPAK was used to summarise and visualise the Q-values. A Q-value threshold of Q-value $\geq$ 0.900 was considered full assignment to either lineage.

Results

Location, health and survival of free-living colonies

Between November 2015 and November 2018, a total of 209 reports of putatively free-living bee colonies were received of which 7.2% (15) were identified as bumble bees, solitary bees, or wasps, and 5.7% (12) were captured swarms of unconfirmed provenance. Colonies reported in early Spring 2016 (prior to expected swarming season), were assumed to have been present since at least late Autumn 2015. The reported habitats of all honey bee colonies (n = 182) consisted primarily of cavities in buildings (68%) where mainly the roof space was occupied by the colony whereas trees formed the second most utilised habitat (10%) (Table 2, Figure 2). Of the 76 colonies that were monitored for survival from Autumn 2015 to Spring 2019, the survival reports on 16 colonies were considered ambiguous (e.g., property containing colony no longer accessible) and removed from further study, 21 colonies (27.63%) survived for 2-2.5 years and 22 (28.95%) survived three or more years (Figure 3). Colony deaths primarily occurred over winter or in early spring. Internal causes of colony death were unknown while external causes of death were extermination by the homeowner and predation by Pine marten (*Martes martes*).

Population structure from mitochondrial data

All 99 COI-COII mitochondrial sequences generated from free-living colonies were designated as *A. m. mellifera* as they contained the P element (Cornuet et al., 1991). After pruning to remove identical sequences, 52 sequences remained; 37 PQQ, 11 PQQQ and four PQQQQ mitotypes (Figure 4). When the entire Q elements were included in the TCS analyses (Clement et al., 2000), unsurprisingly separate networks were formed for sequences that contained different numbers of duplicated Q elements (i.e., PQQ, PQQQ or

Table 2. Reported residency periods (by those who reported them to the authors) and range for 182 honey bee colonies by habitat type.

Habitat	Number (%) of colonies	Mean residency (Years)	Residency range (Years)
Buildings	124 (68%)	4.6	1 to 40
Walls	7 (4%)	6.0	1 to 30
Trees	18 (10%)	5.9	1 to 40
Other	33 (18%)	1.2	1 to 5
All types	**182**	**4.2**	**1 to 40**

"Other" includes, *inter alia*, graves, a statue, a cattle grid and a bird nest box.

PQQQQ). When the Q elements were represented only by one base pair each, plus the seven sites where variation occurred within them, clusters with PQQ, PQQQ or PQQQQ were still evident in the network but the three clusters were now connected (Figure 4) except for a single haplotype (F36C056) and a group of three (F28G267, F35CW055 and F51MO096) which were distinct.

Trimming of the 5′ and 3′ ends of the alignments led to no differentiation between the M4d and M4e, and the M4a and M4m, haplotypes. As shown in Figure 4, just over half of colonies (27, 52%) yielded sequences that were identical to M4e and M4d reference haplotypes and two of the Irish free-living bee sequences were identical to M4f, while the rest of the mitochondrial sequences (23, 44%) were distinct from any available European sequence in GenBank. In a comparison of the data from free-living and managed bees sampled to date from Ireland, seven variants were found only in the free-living bees. Four of these sequences (F31G50, F45G085, F13R036, and F43G080/FKe017) had variations represented by single indels or point mutations. Three variants (F6L011, F36C056 and F28G267/F35CW055/F51MO096) had multi-base deletions of between 6 bp and 10 bp (Figure 4). A fourth variant (F18G104/F47G256), with a significant deletion (6 bp from sites 17 to 22), was found in the Irish managed population but not in any available data from elsewhere in Europe.

Subspecies purity of free-living bees from microsatellite and SNP data

Both microsatellite and SNP data indicate a clear structure between the C-lineage bees and the M-lineage bees for K = 2 with Irish free-living bees showing a high degree of purity with both sets of markers (Table 1 and Figure 5 (for the microsatellite data)). The free-living colony F21CE (Table 1) was located in County Clare in an area known to contain beekeepers that keep 'Buckfast'. This colony contained a mixture of bees with different levels of putative purity, with some bees being assigned to M lineage with Q-value of 0.998 while one bee had a Q-value of 0.607. Colony F44G was collected in the walls of a castle in east Galway and again showed bees with mixed ancestry, some assigned to M lineage with Q-value of 0.997 and others dropping below the 0.900, cut off for confident lineage assignment. Similarly, colony F55MN, collected in the roof of an old cottage in County Monaghan, showed one bee that could be assigned to M lineage with confidence while the other bee could not (Table 1).

SNP data identified only one colony falling below a Q-value threshold of 0.900, and this was colony F21CE, also identified by microsatellites as showing introgression (Table 1). While lineage assignment was not tested for colony F44G via the SNP approach, the bee from colony F55MN had a Q-value of 0.967. Where SNPs data was returned for 8/9 bees per colony, all bees from all 12 tested colonies could be clearly assigned to the M lineage. Unfortunately, colonies F21CE, F44G and F55MN were not included in that experiment.

AMOVA (carried out including only one bee per colony) indicated segregation of the M and C lineages with 33.28% of the genetic variation evident between groups. The greatest proportion of variation occurred within the populations (63.76%), with little distinction between populations within groups. A low pairwise F_{ST} value (0.001), indicated little distinction between the managed and free-living honey bee cohorts in Ireland. Average gene diversity (0.53 ± 0.28) was the same in the two groups of bees whilst observed heterozygosity was slightly lower in the managed honey bee cohort (0.47 compared with 0.51). The numbers of alleles found across populations were not fully comparable due to the differences in population sizes included. While numbers of alleles from the 64 managed bees (128) was higher than the 113 alleles reported from the 54 free-living bees both numbers are far lower than the 240 alleles reported from the 182 *A. m. carnica* bees from Austria but higher than the numbers of alleles returned from all the other populations included.

Pairwise Fst values between the free-living Irish cohort and the *A. m. mellifera* populations from Switzerland and Norway (0.045 and 0.067 respectively) was an order of magnitude lower than pairwise values between Irish bees and those from Sweden and France (0.12 and 0.0.18 respectively). Fst values between the Irish bees and *A. m. ligustica* and *A. m. carnica* were between 0.32 and 0.55 reflecting the subspecies differentiation. Genetic diversity in the Irish free-living cohort was similar but slightly higher than the *A. m. ligustica* population sampled from Italy with similar numbers of bees included in analyses (0.53 versus 0.31). Average gene diversity estimates of the other *A. m. mellifera* populations were similar to those from Ireland (ranging

Figure 2. Images of some locations from where free-living honey bee colonies were collected. A. cavity in a tree showing honey bees at entrance and propolis staining, B. Sampling bees from a house roof fascia with a bee vacuum, C. holes in castle walls are a common location for free-living honey bee colonies (the location of two colonies are indicated by the arrows), D. Tree that hosts a colony of free-living honey bees for over four years (the arrow points to the colony entrance).

from 0.37 France to 0.56 Sweden) despite the higher numbers of bees included in the Irish sample set.

Discussion

This study provides the first fully documented evidence, since the discovery of *V. destructor* in Ireland in 1998, that free-living honey bees exist in the country, primarily using the cavities provided in new and historic buildings to house their colonies. While the majority of colonies have been monitored as part of this study for only three years, some of these colonies had already been in place for a period prior to the study commencing. Very few sampled free-living colonies showed signs of introgression from introduced subspecies and hybrids and we provide sequences of novel mitochondrial haplotypes from this cohort.

In contrast to our findings where 68% of colonies were located in buildings, in the UK there was no significant difference between the use of trees, houses and walls for colony sites (Thompson, 2012). In Ireland approximately 11% land cover is woodland, the majority of which is commercially grown coniferous species (D.A.F.M., 2018). A high turnover of trees in managed forests combined with loss of mature deciduous woodland is likely to produce a low density of trees with cavities of sufficient size for colonisation by honey bees. Conversely, although the UK has a

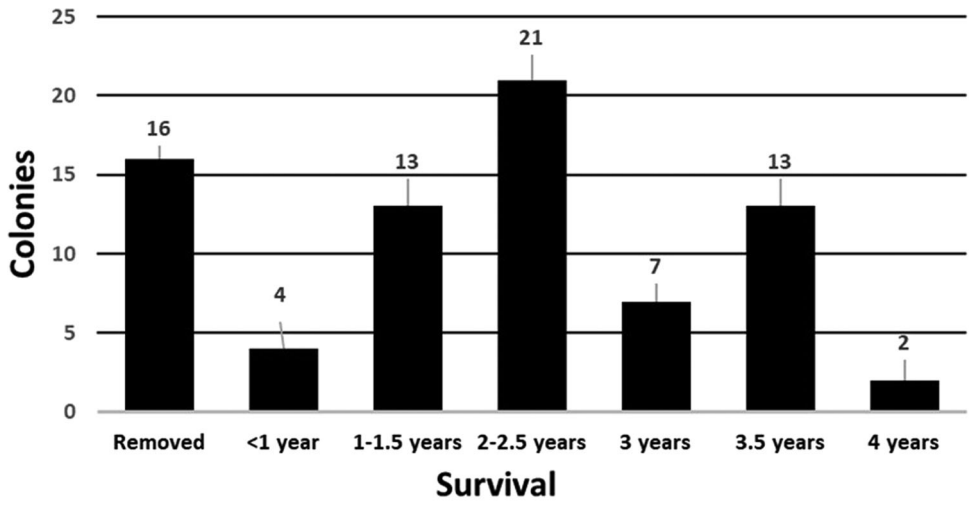


Figure 3. Survival of colonies (n = 76) monitored between Autumn 2015 and Spring 2019. Sixteen colonies were removed from the study due to incomplete records or ambiguity regarding survival as reported by the custodian, 22 have survived at least three winters of which two have survived four. In some cases, custodians reported that locations have been housing colonies for decades.

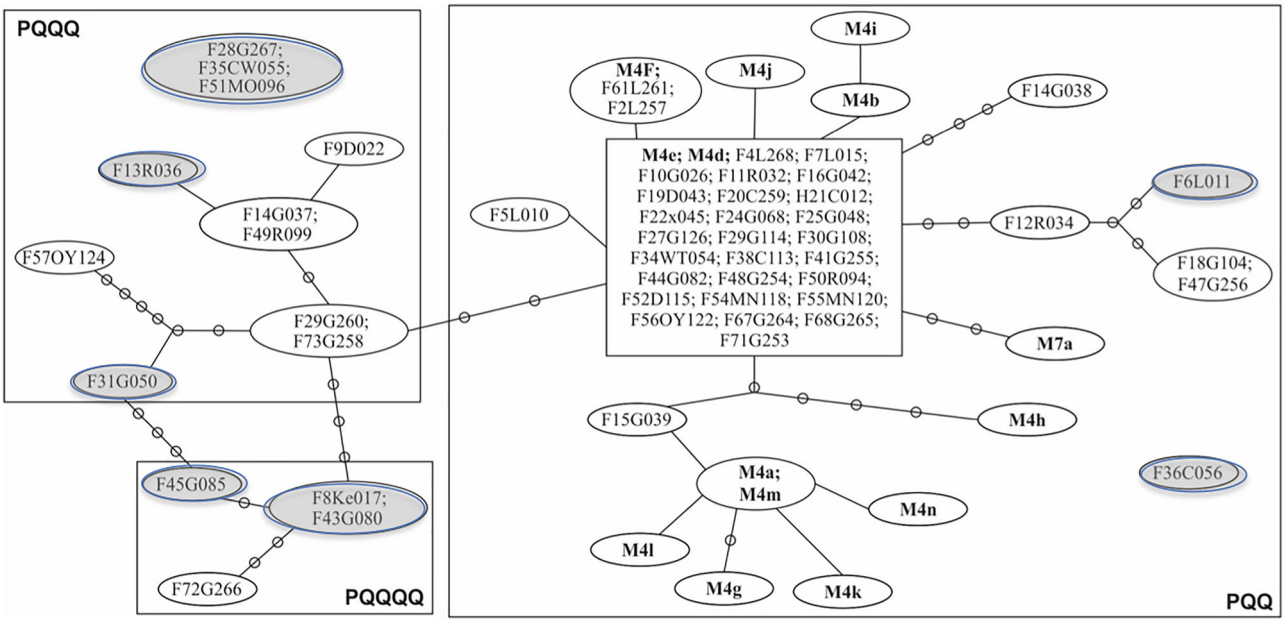


Figure 4. Statistical parsimony network of the 19 different mitotypes from Irish free-living bees. These have been included with European haplotypes downloaded from Genbank. Each circle (o) on a branch represents a single indel or point mutation. Branch lengths are not representative of distance. The variants in grey are those variants that were only present in the free-living population in Ireland, being absent from the managed population sampled so far. PQQ: All sequence types in the PQQ box contain the intergenic P motif and two Q repeat elements, PQQQ: all sequences included contain the intergenic P motif and three Q repeat elements, PQQQ: all sequences included contain the intergenic P motif and four Q repeat elements.

similar relative woodland cover (13%), it consists of roughly even areas of coniferous and broadleaved woodland, which creates a higher age profile than that found in Ireland (Forest Research UK, 2018) and this may help explain the relatively greater use of trees by UK colonies. The difference may also be related to sample design with colonies in houses being easier to spot by members of the public. Searches for wild honey bees in old Irish woodland should be made to expand this dataset. However, colonies in house cavities obtain the benefits of a long lasting, insulated space, giving them the time needed to expand both individual colonies and the dynasty of the queen, an arrangement which may be a benefit over the use of tree cavities. It maybe that honey bees have learnt to exploit the increase in building numbers given the decline in other suitable habitat.

As we reported in Hassett et al. (2018) for the Irish honey bee population generally, over 50% of the

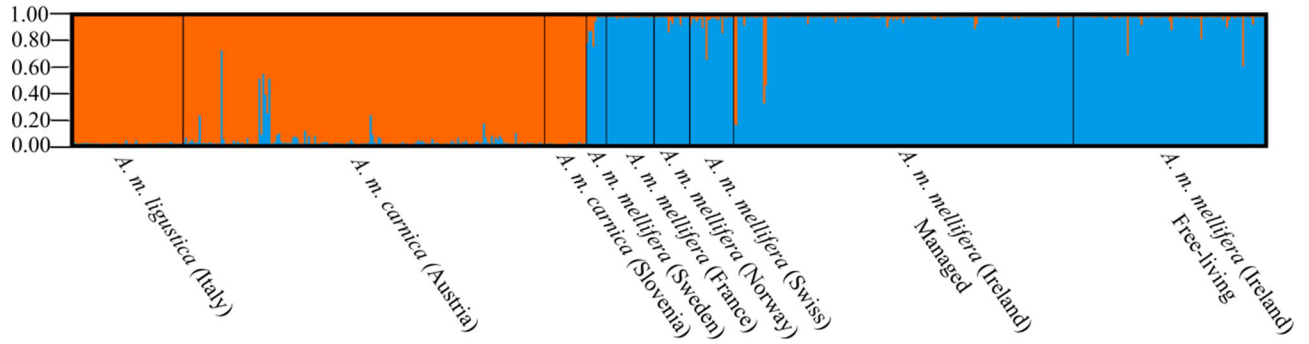


Figure 5. Structure K = 2 assignments of 598 individuals from 9 putative populations (1: Italian *A. m. ligustica*, 2: Austrian *A. m. carnica*, 3: Slovenian *A. m. carnica* 4: Swedish *A. m. mellifera*, 5: French *A. m. mellifera*, 6: Norwegian *A. m. mellifera*, 7: Swiss *mellifera* 8: Irish managed and 9: Irish free-living) including 95 individual honey bees from 50 free-living colonies. Each bar represents an individual bee with assignment apportioned between orange (C-lineage) and blue (M-lineage). Assignment values are from 0.000 to 1.000.

mitochondrial haplotypes identified in the Irish free-living colonies were identical to Dutch haplotypes and are likely descendants from the significant imports that were brought in from Holland. Remarkably none of the free-living bees showed haplotypes from elsewhere in Europe (i.e., not identical to any of the sequences from Rortais et al. (2011) or Pinto et al. (2014)). Hassett et al. (2018) revealed two Irish bees with identical mitochondrial sequences to a French haplotype and one bee showed identity to bees from Colonsay. These bees were from the managed cohort and together the results so far indicate little influence in Ireland of *A. m. mellifera* from mainland Europe, apart from a major influence from the importations from Holland. Most of the mitochondrial variants found in the free-living bees could also be found within the managed Irish population (Hassett et al., 2018), meaning Ireland has a free-living population that appears fundamentally undifferentiated from the managed one, also confirmed by AMOVA of microsatellite data. That said, seven mitochondrial variants were found exclusively in the Irish free-living bees, which might indicate unique genetic variation present in free-living bees. However, this is a relatively small study and additional sampling may identify the putatively exclusive free-living variants in the managed population.

Interestingly, the retention of clusters that are essentially defined by the number of Q elements present may provide some evidence that each duplication occurs as a single event such as the AT-rich homologous motifs suggested by Cornuet et al. (1991). If the duplication event is a synapomorphy, then the Q element architecture represents separate M lineages, creating a problem with the current naming system for mitochondrial haplotypes. Putative mitochondrial lineages as evident in Figure 4 and in Hassett et al. (2018) and more recently in Henriques et al. (2018) will need to be confirmed with other data. Highly significant is that both microsatellite and SNP data indicate that, in keeping with the Irish population as a whole, the free-living population sampled consists mostly of bees that can be assigned to *A. m. mellifera* with high confidence.

Genetic diversity estimates from data generated from the free-living honey bee population in Ireland indicates a healthy population with no evidence of a genetic bottleneck. While results are not directly comparable (given that only some of the same microsatellites were used and were performed via a different laboratory), the levels of heterozygosity and gene diversity reported here are higher than those for France and Sweden reported by Estoup et al. (1995) and for Spain (De la Rúa et al., 2003).

Evidence now hints at colony survival past the stage where Varroosis related death would be expected (Korpela et al., 1992). In periods as short as ten years, experimental survivorship tests in isolated areas have indicated that a balanced host-parasite relationship can develop in colonies from a small population (Fries et al., 2006; Le Conte et al., 2007). The high rate of genetic recombination in honey bees (Beye et al., 2006) may produce a sufficiently diverse population from only a few survivors to allow genes linked to resistance/tolerance mechanisms to quickly proliferate in a small population. Under natural conditions, colony density can return to the levels present before the arrival of Varroa (Mikheyev et al., 2015; Seeley, 2007) and given the low levels of commercialised beekeeping in Ireland, combined with the existence of a putatively large free-living population, it seems plausible that colonies with Varroa resistance/tolerance mechanisms have emerged in Ireland in the 20 years since Varroa was first discovered. Colonies with traits that allow survival in the presence of Varroa probably existed in Irish apiaries, as well as in the wild population, providing the resources on which natural selection could act. The lack of differentiation between the managed and free-living populations sampled so far may not be considered a negative condition if it leads to the discovery of Varroa resistant/tolerant genotypes. Sampling brood from free-living colonies for disease screening is remarkably difficult. Development of methods to access and study pathogens and parasites in resident free-living colonies are now

required to elucidate variability in disease loads and mechanisms of survival/tolerance.

Given the interactions between genotype and environment that have been clearly shown in honey bees (Büchler et al., 2014; Meixner et al., 2014), and the fact that local ecotypes seem to do better than bees translocated from different microclimates (Costa et al., 2012), there is also a need to protect local adaptations. Research on free-living bees in Ireland now requires expansion to allow an accurate indication of colony density on a national scale, together with observation of survivorship and associated mechanisms, as well as characterisation of local strains. We believe that the results presented here, combined with those of Hassett et al. (2018), and the observations of beekeepers, require the immediate application of the precautionary principle of conservation practice (Finnoff et al., 2007), for the protection of Ireland's locally adapted free-living honey bee population. We join with other researchers in requesting legal protection for local adaptations in *A. mellifera* (Fontana et al., 2018) through stricter control on the movement of live bees and the banning of imports. Inter-country movement of *A. m. mellifera,* even where it is the indigenous subspecies, needs to be given careful consideration in each case to avoid out-breeding locally adapted gene complexes. However, having been the welcome recipient of Dutch honey bees following the collapse of Irish beekeeping at the beginning of the 20th century, Ireland may yet be able to return the favour by returning bees of Dutch haplotypes home to the Netherlands from a free-living population.

Acknowledgements

We particularly thank the custodians of the free-living honey bee colonies and the Native Irish Honey Bee Society (NIHBS) for their assistance. KAB is a recipient of an Irish Research Council postgraduate fellowship (GOIPG/2015/2767) and a Tony Ryan Postgraduate fellowship. Additional funding was gratefully received from the Department of Agriculture, Food and the Marine [grant number GRGAS 16/GR/09], the Federation of Irish Beekeeping Associations, the Eva Crane Trust [grant number ECTA20160303] and The Native Irish Honey Bee Society. Financial support for DH was provided through the program COMPETE 2020 – POCI (Programa Operacional para a Competividade e Internacionalização) and by Portuguese funds through FCT (Fundação para a Ciência e a Tecnologia) in the framework of the project BeeHappy (POCI-01-0145-FEDER-029871).

Disclosure statement

The authors declare that there are no conflicts of interest or disputes over ownership of the data and all contributions to this work have been attributed appropriately.

ORCID

Keith A. Browne http://orcid.org/0000-0001-6666-9821
Roberto Ferrari http://orcid.org/0000-0003-4015-8104
Grace P McCormack http://orcid.org/0000-0003-4960-5185

References

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Anderson, D. L., & Trueman, J. W. H. (2000). Varroa jacobsoni (Acari: Varroidae) is more than one species. *Experimental & Applied Acarology*, *24*(3), 165–189. https://doi.org/10.1023/A:1006456720416

Beye, M., Gattermeier, I., Hasselmann, M., Gempe, T., Schioett, M., Baines, J. F., Schlipalius, D., Mougel, F., Emore, C., Rueppell, O., Sirvio, A., Guzman-Novoa, E., Hunt, G., Solignac, M., & Page, R. E. (2006). Exceptionally high levels of recombination across the honey bee genome. *Genome Research*, *16*(11), 1339–1344. https://doi.org/10.1101/gr.5680406

Büchler, R., Costa, C., Hatjina, F., Andonov, S., Meixner, M. D., Conte, Y. L., Uzunov, A., Berg, S., Bienkowska, M., Bouga, M., Drazic, M., Dyrba, W., Kryger, P., Panasiuk, B., Pechhacker, H., Petrov, P., Kezić, N., Korpela, S., & Wilde, J. (2014). The influence of genetic origin and its interaction with environmental effects on the survival of *Apis mellifera* L. colonies in Europe. *Journal of Apicultural Research*, *53*(2), 205–214. https://doi.org/10.3896/IBRA.1.53.2.03

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*(10), 1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x

Cornuet, J. M., Garnery, L., & Solignac, M. (1991). Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA. Genetics, *128*(2), 393–403.

Costa, C., Lodesani, M., & Bienefeld, K. (2012). Differences in colony phenotypes across different origins and locations: Evidence for genotype by environment interactions in the Italian honeybee (*Apis mellifera ligustica*)? Apidologie, *43*(6), 634–642. https://doi.org/10.1007/s13592-012-0138-9

De la Rúa, P., Galián, J., Serrano, J., & Moritz, R. F. A. (2003). Genetic structure of Balearic honeybee populations based on microsatellite polymorphism. *Genetics, Selection, Evolution*, *35*(3), 339–350. https://doi.org/10.1186/1297-9686-35-3-339

De la Rua, P., Jaffe, R., Dall'Olio, R., Munoz, I., & Serrano, J. (2009). Biodiversity, conservation and current threats to European honeybees. Apidologie, *40*(3), 263–284. https://doi.org/10.1051/apido/2009027

D.A.F.M. (2018). (Department of Agriculture Fisheries and Marine) website. Retrieved November 23. www.agriculture.gov.ie/forestservice/

Ellis, J. S., Soland-Reckeweg, G., Buswell, V. G., Huml, J. V., Brown, A., & Knight, M. E. (2018). Introgression in native populations of Apis mellifera mellifera L: Implications for conservation. *Journal of Insect Conservation*, *22*(3-4), 377–390. https://doi.org/10.1007/s10841-018-0067-7

Estoup, A., Garnery, L., Solignac, M., & Cornuet, J.-M. (1995). Microsatellite variation in the honeybee (Apis mellifera L.) populations: Hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics*, *140*, 679–695.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Finnoff, D., Shogren, J. F., Leung, B., & Lodge, D. (2007). Take a risk: Preferring prevention over control of biological invaders. *Ecological Economics*, *62*(2), 216–222. https://doi.org/10.1016/j.ecolecon.2006.03.025

Fontana, P., Costa, C., Di Prisco, G., Ruzzier, E., Annoscia, D., Battisti, A., Caoduro, G., Carpana, E., Contessi, A., Dal Lago, A., Dall 'Olio, R., De Cristofaro, A., Felicioli, A., Floris, I., Fontanesi, L., Gardi, T., Lodesani, M., Malagnini, V., Manias, L., … Segrè, A. (2018). Appeal for biodiversity protection of native honey bee subspecies of *Apis mellifera* in Italy (San Michele all'Adige declaration). *Bulletin of Insectology*, *71*(2), 257–271.

Forest Research UK. (2018). *National forest inventory*. Retrieved November 26, 2018, from www.forestresearch.gov.uk/tools-and-resources/national-forest-inventory/

Fries, I., Imdorf, A., & Rosenkranz, P. (2006). Survival of mite infested (Varroa destructor) honey bee (Apis mellifera) colonies in a Nordic climate. Apidologie, *37*(5), 564–570. https://doi.org/10.1051/apido:2006031

Gabriel, S., Ziaugra, L., & Tabbaa, D. (2009). SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Current Protocols in Human Genetics*, *60*(1), 2.12.1–2.12.18 https://doi.org/10.1002/0471142905.hg0212s60

Garnery, L., Franck, P., Baudry, E., Vautrin, D., Cornuet, J.-M., & Solignac, M. (1998). Genetic diversity of the west European honey bee (Apis mellifera mellifera and A. m. iberica). I. Mitochondrial DNA. *Genetics Selection Evolution*, *30*(Suppl 1), S31–S47. https://doi.org/10.1186/1297-9686-30-S1-S31

Garnery, L., Solignac, M., Celebrano, G., & Cornuet, J. M. (1993). A simple test using restricted PCR-amplified mitochondrial-DNA to study the genetic structure of Aois-mellifera L. Experientia, *49*(11), 1016–1021. https://doi.org/10.1007/BF02125651

Hassett, J., Browne, K. A., McCormack, G. P., Moore, E., Society, N. I. H. B., Soland, G., Geary, M., & Native Irish Honey Bee Society. (2018). A significant pure population of the dark European honey bee (Apis mellifera mellifera) remains in Ireland. *Journal of Apicultural Research*, *57*(3), 337–350. https://doi.org/10.1080/00218839.2018.1433949

Henriques, D., Browne, K. A., Barnett, M. W., Parejo, M., Kryger, P., Freeman, T. C., Munoz, I., Garnery, L., Highet, F., Jonhston, J. S., McCormack, G. P., & Pinto, M. A. (2018). High sample throughput genotyping for estimating C-lineage introgression in the dark honeybee: An accurate and cost-effective SNP-based tool. *Scientific Reports*, *8*(1), 8552. https://doi.org/10.1038/s41598-018-26932-1

Jaffe, R., Dietemann, V., Allsopp, M. H., Costa, C., Crewe, R. M., Dall'olio, R., de la Rua, P., El-Niweiri, M. A. A., Fries, I., Kezic, N., Meusel, M. S., Paxton, R. J., Shaibi, T., Stolle, E., & Moritz, R. F. A. (2010). Estimating the density of honeybee colonies across their natural range to fill the gap in pollinator decline censuses. *Conservation Biology*, *24*(2), 583–593. https://doi.org/10.1111/j.1523-1739.2009.01331.x

Jensen, A. B., Palmer, K. A., Boomsma, J. J., & Pedersen, B. V. (2005). Varying degrees of Apis mellifera ligustica introgression in protected populations of the black honeybee, Apis mellifera mellifera, in northwest Europe. *Molecular Ecology*, *14*(1), 93–106. https://doi.org/10.1111/j.1365-294X.2004.02399.x

Kohl, P., & Rutschmann, B. (2018). The neglected bee trees: European beech forests as a home for feral honey bee colonies. *PeerJ*, *6*(3), e4602.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Korpela, S., Aarhus, A., Fries, I., & Hansen, H. (1992). Varroa jacobsoni OUD in cold climates: Population growth, winter mortality and influence on survival of honey bee colonies. *Journal of Apicultural Research*, *31*(3-4), 157–164. https://doi.org/10.1080/00218839.1992.11101278

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Le Conte, Y., De Vaublanc, G., Crauser, D., Jeanne, F., Rousselle, J.-C., & Becard, J.-M. (2007). Honey bee colonies that have survived Varroa destructor. Apidologie, *38*(6), 566–572. https://doi.org/10.1051/apido:2007040

Meixner, M. D., Büchler, R., Costa, C., Francis, R. M., Hatjina, F., Kryger, P., Uzunov, A., & Carreck, N. L. (2014). Honey bee genotypes and the environment. *Journal of Apicultural Research*, *53*(2), 183–187. https://doi.org/10.3896/IBRA.1.53.2.01

Mikheyev, A. S., Tin, M. M. Y., Arora, J., & Seeley, T. D. (2015). Museum samples reveal rapid evolution by wild honey bees exposed to a novel parasite. *Nature Communications*, *6*(1), 7991. https://doi.org/10.1038/ncomms8991

Moritz, R. F. A., Bernhard Kraus, F., Kryger, P., & Crewe, R. M. (2007). The size of wild honeybee populations (Apis mellifera) and its implications for the conservation of honeybees. *Journal of Insect Conservation*, *11*(4), 391–397. https://doi.org/10.1007/s10841-006-9054-5

Moritz, R. F. A., Hartel, S., & Neumann, P. (2005). Global invasions of the western honeybee (Apis mellifera) and the consequences for biodiversity. Écoscience, *12*(3), 289–301. https://doi.org/10.2980/i1195-6860-12-3-289.1

Nieto, A., Roberts, S. P. M., Kemp, J., Rasmont, P., Kuhlmann, M., García Criado, M., Biesmeijer, J. C., Bogusch, P., Dathe, H. H., De la Rúa, P., De Meulemeester, T., Dehon, M., Dewulf, A., Ortiz-Sánchez, F. J., Lhomme, P., Pauly, A., Potts, S. G., Praz, C., Quaranta, M., … Michez, D. (2014). *European Red List of bees*. Publication Office of the European Union.

Parejo, M., Henriques, D., Pinto, M. A., Soland-Reckeweg, G., & Neuditschko, M. (2018). Empirical comparison of microsatellite and SNP markers to estimate introgression in Apis mellifera mellifera. *Journal of Apicultural Research*, *57*(4), 504–506. https://doi.org/10.1080/00218839.2018.1494894

Pinto, M. A., Henriques, D., Chávez-Galarza, J., Kryger, P., Garnery, L., van der Zee, R., Dahle, B., Soland-Reckeweg, G., de la Rúa, P., Dall' Olio, R., Carreck, N. L., & Johnston, J. S. (2014). Genetic integrity of the Dark European honey bee (Apis mellifera mellifera) from protected populations: A genome-wide assessment using SNPs and mtDNA sequence data. *Journal of Apicultural Research*, *53*(2), 269–278. https://doi.org/10.3896/IBRA.1.53.2.08

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology*, *17*(1), 285–293. https://doi.org/10.1111/j.1365-294X.2007.03417.x

Requier, F., Garnery, L., Kohl, P. L., Njovu, H. K., Pirk, C. W. W., Crewe, R. M., & Steffan-Dewenter, I. (2019). The conservation of native honey bees is crucial. *Trends in Ecology & Evolution*, *34*(9), 789–798. https://doi.org/10.1016/j.tree.2019.04.008

Rortais, A., Arnold, G., Alburaki, M., Legout, H., & Garnery, L. (2011). Review of the Dral COI-COII test for the conservation of the black honeybee (Apis mellifera mellifera). *Conservation Genetics Resources*, *3*(2), 383–391. https://doi.org/10.1007/s12686-010-9351-x

Seeley, T. D. (2007). Honey bees of the Arnot Forest: A population of feral colonies persisting with Varroa destructor in the northeastern United States. Apidologie, *38*(1), 19–29. https://doi.org/10.1051/apido:2006055

Soland-Reckeweg, G. (2006). *Genetic differentiation and hybridization in the honeybee (Apis mellifera L.) in Switzerland* [Thesis]. University of Bern.

Soland-Reckeweg, G., Heckel, G., Neumann, P., Fluri, P., & Excoffier, L. (2009). Gene flow in admixed populations and implications for the conservation of the Western honey bee. *Journal of Insect Conservation*, *13*(3), 317–328. https://doi.org/10.1007/s10841-008-9175-0

Thompson, C. E. (2012). *The health and status of the feral honeybee (Apis mellifera sp) and the Apis mellifera mellifera population of the UK* [Thesis]. University of Leeds.

Vaha, J. P., & Primmer, C. R. (2006). Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, *15*(1), 63–72. https://doi.org/10.1111/j.1365-294X.2005.02773.x