<table>
<thead>
<tr>
<th>Title</th>
<th>Functional Analysis of Vertebrate Msl2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Lai, Zheng</td>
</tr>
<tr>
<td>Publication Date</td>
<td>2014-02-12</td>
</tr>
<tr>
<td>Item record</td>
<td><a href="http://hdl.handle.net/10379/4466">http://hdl.handle.net/10379/4466</a></td>
</tr>
</tbody>
</table>

Some rights reserved. For more information, please see the item record link above.
Functional Analysis of Vertebrate Msl2

Zheng Lai

Epigenetics and Chromatin Biology Laboratory,
Centre for Chromosome Biology,
School of Natural Sciences,
National University of Ireland, Galway

A thesis submitted to the National University of Ireland, Galway
for the degree of Doctor of Philosophy

September 2013

Supervisor: Dr. Stephen Rea
# Table of contents

Table of contents ........................................................................................................i
List of Figures ........................................................................................................... iv
List of Tables ............................................................................................................ vi
Acknowledgements .................................................................................................. vii
Declaration ............................................................................................................... viii
Abbreviations .......................................................................................................... ix
Abstract .................................................................................................................. xiii

## Chapter 1 Introduction

1.1 Chromatin structure and histone modifications ................................................. 1
   1.1.1 Chromatin structure .................................................................................... 1
   1.1.2 Histone modifications ................................................................................ 2
      1.1.2.1 Histone acetylation ............................................................................. 4
      1.1.2.2 Histone phosphorylation ................................................................... 6
      1.1.2.3 Histone ubiquitination ....................................................................... 7
1.2 Ubiquitination and RING finger domain E3 ligases ........................................ 8
   1.2.1 Ubiquitination ............................................................................................ 8
   1.2.2 RING domain E3 ligases .......................................................................... 11
1.3 The MSL complex ............................................................................................ 12
   1.3.1 The MSL complex in Drosophila ............................................................... 12
   1.3.2 The MSL complex in Human .................................................................. 14
      1.3.2.1 Human MOF ...................................................................................... 14
      1.3.2.2 Human MSL1 ................................................................................... 17
      1.3.2.3 Human MSL2 ................................................................................... 17
1.4 DNA damage ..................................................................................................... 18
   1.4.1 DNA damage response .............................................................................. 18
   1.4.2 DNA damage repair .................................................................................. 21
      1.4.2.1 Homologous recombination repair pathway (HR) ......................... 24
      1.4.2.2 Non-homologous end joining pathway (NHEJ) .............................. 24
      1.4.2.3 Microhomology-mediated end joining (MMEJ or alt-NHEJ) ...... 25
   1.4.3 Histone modifications in DNA damage response and repair .......... 26
1.5 Chicken DT40 cell line as a model system ....................................................... 29
   1.5.1 Chicken DT40 cell line ............................................................................ 29
   1.5.2 Gene targeting in DT40 cell line .............................................................. 30
1.6 Aim of the thesis .............................................................................................. 31

## Chapter 2 Materials and methods

2.1 Materials .......................................................................................................... 32
   2.1.1 Solutions, buffers and media .................................................................... 32
2.2 Methods .................................................................40
  2.2.1 Nucleic acid methods .........................................40
    2.2.1.1 RNA preparation ...........................................40
    2.2.1.2 Estimation of RNA concentration .....................40
    2.2.1.3 cDNA synthesis ..........................................41
    2.2.1.4 Polymerase chain reaction ............................41
    2.2.1.5 Site directed mutagenesis ...............................42
    2.2.1.6 Generation of probes labelled with Digoxigenin (DIG) by PCR ....42
    2.2.1.7 Quantitative real-time PCR ............................43
    2.2.1.8 Restriction digestion of DNA ............................43
    2.2.1.9 Alkaline phosphatase treatment of DNA ............44
    2.2.1.10 Agarose gel electrophoresis ...........................44
    2.2.1.11 DNA Extraction .........................................44
    2.2.1.12 DNA Ligation ...........................................44
    2.2.1.13 Preparation of competent E. coli by chemical method ....45
    2.2.1.14 E. coli transformation ................................45
    2.2.1.15 Plasmid DNA preparation ..............................45
    2.2.1.16 Estimation of DNA concentration ....................46
    2.2.1.17 DNA sequencing .......................................46
    2.2.1.18 Extraction of genomic DNA from chicken DT40 cells ....46
    2.2.1.19 Southern blot ..........................................47
  2.2.2 Protein methods ................................................48
    2.2.2.1 Protein sample preparation ............................48
      2.2.2.1.1 Whole cell extraction ..............................48
      2.2.2.1.2 Isolation of nuclear and cytoplasmic proteins ....48
    2.2.2.2 Bradford protein assay ..................................48
    2.2.2.3 SDS-polyacrylamide gel electrophoresis ...............49
    2.2.2.4 Coomassie staining .....................................50
    2.2.2.5 Transfer of proteins to membrane .....................50
    2.2.2.6 Immunoblotting .........................................50
  2.2.3 Cell biology techniques ....................................51
    2.2.3.1 Maintenance of cells ..................................51
    2.2.3.2 Freezing and thawing cells ...........................51
    2.2.3.3 Cell proliferation analysis ............................51
Table of contents

2.2.3.4 Cell transfection..........................................................52
2.2.3.4.1 Stable transfection of DT40 cells ..................................52
2.2.3.4.2 Transient transfection of human cells.............................52
2.2.3.5 Clonogenic survival.........................................................53
2.2.3.6 Flow cytometry............................................................53

Chapter 3 Generation of DT40 Msl2\(^{+}\) and Msl2\(^{-}\)-rescue cell lines ......55
3.1 Introduction ...........................................................................55
3.2 Sequence analysis of Msl2 protein ........................................56
3.3 Mapping of the chicken Msl2 genomic locus.............................58
3.4 Cloning chicken Msl2 ..........................................................60
3.5 Generation of DT40 Msl2\(^{+}\) and Msl2\(^{-}\)-rescue cell lines ..............61
   3.5.1 Construction of DT40 Msl2\(^{+}\) cell lines .......................61
   3.5.2 Generation of Msl2\(^{-}\) rescue cell lines ...........................63
3.6 Localization of the chicken Msl2 protein ..................................67
3.7 Discussion.............................................................................68

Chapter 4 Functional analysis of Msl2\(^{+}\) and Msl2\(^{-}\)-rescue DT40 cells......71
4.1 Introduction ...........................................................................71
4.2 Chicken Msl2 is not essential for DT40 viability.......................72
   4.2.1 Cell cycle analysis .........................................................73
4.3 Histone modifications perturbed in Msl2 knockout cells ..........75
4.4 Msl2 involved in DNA damage response ...............................77
   4.4.1 Msl2 is required for HR repair ......................................78
   4.4.2 Msl2 accumulates in response to DNA damage .............79
      4.4.2.1 Msl2 is accumulated by MMS treatment ......................79
      4.4.2.2 Msl2 is stabilised by IR treatment .............................81
4.5 Discussion.............................................................................83

Chapter 5 Conclusions and future perspectives ............................88

References ..................................................................................92

Appendix 1 MDC1 downregulated in hMSL1-depleted cells ..............108
Appendix 2 Detection of chicken Msl2 and Mof expression by human
      antibodies .............................................................................109
Appendix 3 Partial length chicken Mof sequence ..........................116
Appendix 4 Scientific communications ........................................117
List of Figures

| Figure 1.1 | Organisation of Chromatin in the eukaryotic nucleus. | 2 |
| Figure 1.2 | Post-translational modifications of human core histones | 3 |
| Figure 1.3 | Ubiquitination cascade | 9 |
| Figure 1.4 | The Drosophila MSL complex | 12 |
| Figure 1.5 | The DNA damage response pathway | 19 |
| Figure 1.6 | DNA repair pathways involved in the repair of double-strand breaks | 23 |
| Figure 3.1 | Analysis of Msl2 protein sequences from human, mouse, chicken, Xenopus and Drosophila. | 57 |
| Figure 3.2 | Multiple alignment of the sequence of Msl2/hMSL2 between chicken and human species | 58 |
| Figure 3.3 | Syntenic gene profile of Msl2 loci from chicken, human and mouse | 59 |
| Figure 3.4 | Cloning of chicken Msl2 cDNA | 60 |
| Figure 3.5 | Strategy for creation of DT40 Msl2 knockout cell line | 61 |
| Figure 3.6 | Testing of the nonradiolabelled 5’ probe by Southern Blot hybridization | 62 |
| Figure 3.7 | Identification of Msl2 knockout cell lines by Southern Blot and RT-PCR | 63 |
| Figure 3.8 | Expression of full length or mutant Msl2 in U2OS cells | 64 |
| Figure 3.9 | Screening stable Msl2 rescue DT40 clones | 65 |
| Figure 3.10 | Screening truncated or point mutated Msl2 re-expressed stable DT40 clones | 67 |
| Figure 3.11 | Nuclear localization of chicken Msl2 protein in DT40 cells | 68 |
| Figure 4.1 | Proliferation analysis of Wild-type, Msl2−/−, Msl2-rescue and Msl2-ΔR cell lines | 73 |
| Figure 4.2 | Flow cytometry analysis of cell cycle progression of asynchronous wild-type, Msl2−/− and Msl2-rescue cells | 74 |
| Figure 4.3 | Histone modifications are perturbed by loss of Msl2 | 76 |
| Figure 4.4 | Survival curves of DT40 cells in methylcellulose colony assay | 79 |
| Figure 4.5 | Msl2 is accumulated following MMS-induced DNA damage | 80 |
| Figure 4.6 | Msl2 is stabilized following DNA damage | 82 |
| Figure 4.7 | Both Mutant Msl2 proteins are stabilized following DNA damage | 83 |
| Figure 5.1 | Schematic representation of the potential function of Msl2 suggested in DNA damage response | 90 |
| Figure A1.1 | MDC1 downregulated in hMSL1-depleted cells | 108 |
| Figure A2.1 | Testing of commercial anti-hMSL2 antibodies in DT40 cells | 110 |
| Figure A2.2 | Testing of monoclonal anti-hMSL2 antibody in DT40 cells | 111 |
| Figure A2.3 | Testing of monoclonal anti-hMSL2 antibody in Msl2 transfected U2OS cells | 112 |
List of Figures

Figure A2.4 Testing of the anti-hMSL2 antibody 8D2 in Msl2 transfected U2OS cells ................................................................. 113
Figure A2.5 Alignment of the assembled chicken Mof sequence and hMOF. . 114
Figure A2.6 Testing anti-hMOF antibody supernatants in DT40 cells .......... 115
List of Tables

Table 2.1 Common reagents and buffers.................................................................32  
Table 2.2 Molecular biology kits in this study ....................................................34  
Table 2.3 Molecular biology kits in this study Primers.........................................34  
Table 2.4 List of vectors and plasmids .................................................................36  
Table 2.5 Strains of E. coli used in this study .......................................................36  
Table 2.6 DT40 cell lines used in this study .........................................................37  
Table 2.7 Mammalian cell lines used in this study ...............................................37  
Table 2.8 Drugs used for stable cell line selection ...............................................38  
Table 2.9 List of siRNA sequences .....................................................................38  
Table 2.10 Primary antibodies used in this study ...............................................39  
Table 2.11 Secondary antibodies used in this study .............................................40  
Table 2.12 Reverse transcription reaction conditions .........................................41  
Table 2.13 PCR reaction conditions for KOD polymerase .................................41  
Table 2.14 PCR reaction conditions for Taq polymerase ....................................42  
Table 2.15 PCR conditions for generation of probes labelled with DIG ..............42  
Table 2.16 Quantitative real-time PCR conditions ..............................................43  
Table 2.17 Recipe of SDS-polyacrylamide gel ...................................................49  
Table 3.1 Targeting frequencies of the targeting constructs ..................................62
Acknowledgements

I would like to owe my thanks to all the people who ever helping me complete this work. Without all of you, I can’t image this coming true to me.

First, I would like to thank my supervisor Dr. Stephen Rea for giving me this treasure opportunity to undertake my PhD studies in his lab and for his scientific guidance, encouragement, and patience in the last 4 years. Also, I am really grateful to have a chance to enjoy a different culture in Ireland. Thank you so much.

I would also like to thank Dr. Sara Buonomo, Dr. Noel F. Lowndes and Dr. Bob Lahue for agreeing to be on my viva committee.

I really appreciate my labmates Jenn, Sandra, Simona, Lukasz, and Anna for their great company and invaluable friendship in the lab. You guys are the best.

Thanks for all your help. I would never forget you wherever I am.

Special thanks to Anne-Marie and Patricia for their assistance with correction of the thesis, to Loretta, Laura and Indu for their supporting in the writing up room.

Many thanks to all the rest of members past and present from the Centre for Chromosome Biology (CCB) for all their timely help and advice.

I also don’t want to miss the opportunity to thank SFI and Beckman Funding for the financial assistance to complete my PhD.

I am deeply grateful to my parents’ constant support. Without their incredible love and affection, I won’t be who I am today. I love you, Mum and Dad.

Finally, I want to thank Yifan. Thank you for being my Husband and always being attentive and supportive. It is a wonderful and unforgettable experience that we support each other for the same goal during the last several years. Love you and see you soon.

Zheng
Declaration

I hereby declare that all the work presented here was performed by me without any unauthorised source of assistance except the work shown in Figure 4.3 was done in collaboration with Lukasz P. Andrzejewski and Dervla M. Walshe. I hereby declare that no any similar version of this thesis has been submitted as a dissertation to any other university or institute.

Zheng Lai

30th September 2013
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9-1-1</td>
<td>RAD9–RAD1–HUS1</td>
</tr>
<tr>
<td>53BP1</td>
<td>p53-binding protein 1</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine diphosphate</td>
</tr>
<tr>
<td>ALLN</td>
<td>N-Acetyl-L-leucyl-L-leucyl-L-norleucinal</td>
</tr>
<tr>
<td>alt-NHEJ</td>
<td>Alternative NHEJ</td>
</tr>
<tr>
<td>AT</td>
<td>Ataxia telangiectasia</td>
</tr>
<tr>
<td>ATCC</td>
<td>American tissue culture collection</td>
</tr>
<tr>
<td>ATM</td>
<td>Ataxia telangiectasia mutated</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>ATR</td>
<td>Ataxia telangiectasia and Rad3 related</td>
</tr>
<tr>
<td>BARD1</td>
<td>BRCA1-associated RING domain 1</td>
</tr>
<tr>
<td>BER</td>
<td>Base excision repair</td>
</tr>
<tr>
<td>BMI1</td>
<td>B lymphoma Mo-MLV insertion region 1</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Breast cancer 1, early onset</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>c-Cbl</td>
<td>Casitas B-lineage lymphoma</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CHK1/2</td>
<td>Checkpoint kinase 1/2</td>
</tr>
<tr>
<td>COMPASS</td>
<td>Complex proteins associated with Set1</td>
</tr>
<tr>
<td>CtIP</td>
<td>C-terminal binding protein interacting protein</td>
</tr>
<tr>
<td>Ctk1</td>
<td>Catalytic (alpha) subunit of C-terminal domain kinase I</td>
</tr>
<tr>
<td>CXC</td>
<td>Cysteine-rich</td>
</tr>
<tr>
<td>DAPI</td>
<td>4′,6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>DDR</td>
<td>DNA damage response</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s modified Eagle’s medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA-PK</td>
<td>DNA-dependent protein kinase</td>
</tr>
<tr>
<td>DNA-PKcs</td>
<td>Catalytic subunit of DNA-PK</td>
</tr>
<tr>
<td>DOT1</td>
<td>Disruptor of telomeric silencing</td>
</tr>
<tr>
<td>DSB</td>
<td>Double strand break</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>DUB</td>
<td>Deubiquitinase</td>
</tr>
<tr>
<td>E1</td>
<td>Ubiquitin activating enzyme</td>
</tr>
<tr>
<td>E2</td>
<td>Ubiquitin conjugating enzyme</td>
</tr>
<tr>
<td>E3</td>
<td>Ubiquitin ligase enzyme</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GCN5</td>
<td>General control of amino acid synthesis protein 5-like 2</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>GNAT</td>
<td>GCN5-related N-acetyltransferase</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>Gy</td>
<td>Gray</td>
</tr>
<tr>
<td>H2AK119ub</td>
<td>Histone H2A lysine 119 monoubiquitination</td>
</tr>
<tr>
<td>H2BK34ub</td>
<td>Histone H2B lysine 34 monoubiquitination</td>
</tr>
<tr>
<td>H2BK120ub</td>
<td>Histone H2B lysine 120 monoubiquitination</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>Histone 3 lysine 4 trimethylation</td>
</tr>
<tr>
<td>H3K9me2/3</td>
<td>Histone 3 lysine 9 di/trimethylation</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>Histone 3 lysine 27 trimethylation</td>
</tr>
<tr>
<td>H3K36me3</td>
<td>Histone 3 lysine 36 trimethylation</td>
</tr>
<tr>
<td>H3K79me2</td>
<td>Histone 3 lysine 79 dimethylation</td>
</tr>
<tr>
<td>H3S10ph</td>
<td>Histone 3 serine10 phosphorylation</td>
</tr>
<tr>
<td>H4K16ac</td>
<td>Histone 4 lysine 16 acetylation</td>
</tr>
<tr>
<td>H4K20me2/3</td>
<td>Histone 4 lysine 20 di/trimethylation</td>
</tr>
<tr>
<td>HAT</td>
<td>Histone acetyltransferase</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>HECT</td>
<td>Homologous To E6-Ap Carboxyl Terminus</td>
</tr>
<tr>
<td>HECT E3</td>
<td>HECT domain E3 ligase</td>
</tr>
<tr>
<td>hMOF</td>
<td>Human MOF</td>
</tr>
<tr>
<td>hMSL1/2</td>
<td>Human MSL1/2</td>
</tr>
<tr>
<td>HMT</td>
<td>Histone methyltransferase</td>
</tr>
<tr>
<td>HOXA9</td>
<td>Homeobox protein HOX-A9</td>
</tr>
<tr>
<td>HR</td>
<td>Homologous recombination</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>IR</td>
<td>Ionizing radiation</td>
</tr>
<tr>
<td>IRIIF</td>
<td>IR-induced nuclear foci</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>K-HMT</td>
<td>Lysine-specific HMT</td>
</tr>
<tr>
<td>Ku70/80</td>
<td>Lupus Ku autoantigen protein p70/80</td>
</tr>
<tr>
<td>LAP</td>
<td>Leukemia associated protein</td>
</tr>
<tr>
<td>LB Broth</td>
<td>Lysogeny broth</td>
</tr>
<tr>
<td>LSD1</td>
<td>Lysine specific demethylase 1</td>
</tr>
<tr>
<td>MCRS2</td>
<td>Microspherule protein 2</td>
</tr>
<tr>
<td>MDC1</td>
<td>Mediator of DNA damage checkpoint 1</td>
</tr>
<tr>
<td>MDM2</td>
<td>Murine double minute 2</td>
</tr>
<tr>
<td>MEIS1</td>
<td>Myeloid ecotropic viral integration site 1</td>
</tr>
<tr>
<td>MEKK1</td>
<td>Mitogen-activated protein kinase kinase</td>
</tr>
<tr>
<td>MLE</td>
<td>Maleless</td>
</tr>
<tr>
<td>MLL1</td>
<td>Mixed lineage in leukemia 1</td>
</tr>
<tr>
<td>MMEJ</td>
<td>Microhomology-mediated end joining</td>
</tr>
<tr>
<td>MMR</td>
<td>Mismatch repair</td>
</tr>
<tr>
<td>MMS</td>
<td>Methylmethane sulfonate</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MSL</td>
<td>Male specific lethal</td>
</tr>
<tr>
<td>MOF</td>
<td>Males absent on the first</td>
</tr>
<tr>
<td>MRN</td>
<td>MRE11/RAD50/NBS1</td>
</tr>
<tr>
<td>MYST</td>
<td>Moz-Ybf2/Sas3-Sas2-Tip60</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>ncRNA</td>
<td>Noncoding RNA</td>
</tr>
<tr>
<td>NER</td>
<td>Nucleotide excision repair</td>
</tr>
<tr>
<td>NHEJ</td>
<td>Non-homologous end joining</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localization sequence</td>
</tr>
<tr>
<td>NSCLC</td>
<td>Non-small cell lung cancer</td>
</tr>
<tr>
<td>NSL</td>
<td>Non-specific lethal</td>
</tr>
<tr>
<td>Nup153</td>
<td>Nucleoporin 153</td>
</tr>
<tr>
<td>p53</td>
<td>Tumor protein 53</td>
</tr>
<tr>
<td>PARP1/2</td>
<td>Poly(ADP-ribose) polymerase 1/2</td>
</tr>
<tr>
<td>PHD</td>
<td>Plant homeodomain</td>
</tr>
<tr>
<td>PIKK</td>
<td>Phosphatidylinositol 3-kinase-like protein kinase</td>
</tr>
<tr>
<td>PNK</td>
<td>Polynucleotide kinase</td>
</tr>
<tr>
<td>PP2A</td>
<td>Phosphatase 2A</td>
</tr>
<tr>
<td>pRNA</td>
<td>Promoter-associated RNA</td>
</tr>
<tr>
<td>PRC1</td>
<td>Polycomb repressive complex 1</td>
</tr>
<tr>
<td>PTIP</td>
<td>Pax transactivation-domain interacting protein</td>
</tr>
<tr>
<td>RAD18</td>
<td>Ring finger protein 73</td>
</tr>
<tr>
<td>Rad51</td>
<td>Radiation sensitive</td>
</tr>
<tr>
<td>RCC</td>
<td>Renal cell carcinoma</td>
</tr>
<tr>
<td>R-HMT</td>
<td>Arginine-specific HMT</td>
</tr>
<tr>
<td>RING</td>
<td>Really interesting new gene</td>
</tr>
<tr>
<td>RING E3</td>
<td>RING domain E3 ligase</td>
</tr>
<tr>
<td>RNF2</td>
<td>Ring finger protein 2</td>
</tr>
<tr>
<td>RNF8</td>
<td>Ring finger protein 8</td>
</tr>
<tr>
<td>RNF168</td>
<td>Ring finger protein 168</td>
</tr>
<tr>
<td>RNF20/40</td>
<td>Ring finger protein 20/40</td>
</tr>
<tr>
<td>RPA</td>
<td>Replication protein A</td>
</tr>
<tr>
<td>roX</td>
<td>RNA on X</td>
</tr>
<tr>
<td>RTK</td>
<td>Receptor protein tyrosine kinase</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>snRNP</td>
<td>Small nuclear ribonucleoprotein</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SIRT1</td>
<td>Sirtuin-1</td>
</tr>
<tr>
<td>SSB</td>
<td>Single strand break</td>
</tr>
<tr>
<td>SSBR</td>
<td>Single strand break repair</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stranded DNA</td>
</tr>
<tr>
<td>SXL</td>
<td>Sex lethal protein</td>
</tr>
<tr>
<td>TIP60</td>
<td>Tat-interactive protein</td>
</tr>
<tr>
<td>TopBP1</td>
<td>Topoisomerase II-binding protein 1</td>
</tr>
</tbody>
</table>
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ub</td>
<td>Ubiquitin</td>
</tr>
<tr>
<td>Ubc13</td>
<td>Ubiquitin-conjugating 13</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>V(D)J</td>
<td>Variable diversity joining</td>
</tr>
<tr>
<td>Wip1</td>
<td>Wild type p53-induced phosphatase</td>
</tr>
<tr>
<td>XRCC1/4</td>
<td>X-ray repair cross-complementing protein 1/4</td>
</tr>
</tbody>
</table>
Abstract

hMSL2 (Male Specific Lethal 2, human) is a RING finger protein with E3 ubiquitin ligase activity. Although it has been shown to target histone H2B at lysine 34 and p53 at lysine 351, suggesting roles in transcription regulation and apoptosis, its molecular contribution to these and other processes remains poorly defined. To better understand the function of this protein in vertebrates we decided to take a knockout approach using the DT40 model system.

We first confirmed the expression of Msl2 in chicken DT40 cells, designed a strategy to target this gene and then confirmed its successful disruption. Msl2−/− cells are viable, with minor growth retardation and mitotic defects. Biochemical analysis of the chromatin in these cells revealed aberrations in the levels of several histone modifications involved in DNA damage response pathways. To ensure that any phenotypes observed in the Msl2−/− cells were due to loss of Msl2 and more specifically, due to its E3 ligase activity, Msl2-rescue, Msl2-ΔR and Msl2-C44R cell lines were made to re-express full length wild-type Msl2, 1-107 amino acids truncated Msl2 or Msl2 with a point mutation C44R respectively. We found that the above defects could be rescued in the Msl2-rescue cell line. Interestingly, when treated with either of the damaging agents ionizing radiation (IR) or methyl-methanesulfonate (MMS), the Msl2-rescue cell line showed increased Msl2 protein stability.

In summary, these data identify a novel role for Msl2 in the cellular response to DNA damage. Moreover, Msl2 plays a role in maintaining normal histone modification profiles, which may also contribute to the DNA damage response.
Chapter 1 Introduction

1.1 Chromatin structure and histone modifications

1.1.1 Chromatin structure

In eukaryotic cells, the genome is assembled as chromatin, which provides a dynamic platform facilitating all DNA-templated processes such as transcription, replication and repair. Chromatin is composed of arrays of repeated nucleosome units. Each nucleosome is composed of 147 bp of DNA wrapped 1.7 times around a histone octamer containing two copies of four core histone proteins; H2A, H2B, H3, and H4 (Kornberg and Lorch, 1999). Together with linker histones and protein scaffolds, these nucleosomes fold and wrap the chromatin into higher order structure to package the genome into the nucleus (Figure 1.1).

Generally, chromatin architecture can be altered in three different ways. The first is to replace the canonical histones with different histone variants (Talbert and Henikoff, 2010), the second is to reposition or evict histones from DNA by ATP-dependent chromatin remodeling (Hargreaves and Crabtree, 2011) and the third way is to covalently modify histones by post-translational modifications (Zentner and Henikoff, 2013). These changes in chromatin organization facilitate the regulation of DNA-templated processes either by changing the accessibility of the chromatin, or by recruiting protein complexes with specific activities (Jenuwein and Allis, 2001).
Figure 1.1 Organisation of chromatin in the eukaryotic nucleus. Schematic depicting the organized packaging of DNA into higher order chromatin to facilitate its compaction into the nucleus. (Pearson 2011).

1.1.2 Histone modifications

The core histones are small, positively charged proteins that are highly conserved during evolution. They are predominantly globular, while their N-terminal 20-35 basic amino acid-rich residue tails are unstructured and highly dynamic (Kouzarides, 2007; Munshi et al., 2009). These tails can protrude from the surface of the nucleosome and are subject to many covalent modifications (Figure 1.2), which play an important role in regulation of higher order chromatin structure and biological processes.
Histone modifications are receiving more and more attention in chromatin research. In the past decades, more than 100 distinct histone modifications were discovered (Zentner and Henikoff, 2013). Among them, at least eight distinct types of histone modifications have been elucidated, including acetylation, phosphorylation, methylation, ubiquitination, sumoylation, deamination poly(ADP-ribosyl)ation, and proline isomerization (Kouzarides, 2007). These modifications happen at many sites on histones. With specific antibodies or mass spectrometry, more than 60 different histone residues have been detected to be modified. Due to the limitation of the current detection methods, the extent of histone modifications is presumed to be much underestimated.

Nowadays, it is generally accepted that histone modifications function in the following two ways. One is to affect the high order structure of chromatin by disruption of contacts of histones in adjacent nucleosomes or interactions of
histones and DNA. The other is the recruitment (or the prevention thereof) of non-histone proteins onto chromatin (Kouzarides, 2007).

Histone modifications are highly dynamic and reversible. It has been found that some of them can very quickly be established and removed on chromatin in response to environmental stimuli. Their half-lives are probably only a few minutes (Barth and Imhof, 2010).

Histone modifications such as acetylation, methylation, phosphorylation and ubiquitination have been implicated in fundamental cellular processes such as epigenetic regulation of gene expression, organization of chromatin structure, chromosome segregation, DNA replication and DNA repair. In the following sections, we will discuss some of these modifications that are relevant to this thesis in more detail.

1.1.2.1 Histone acetylation

Of all the histone modifications, acetylation discovered in 1961 (Phillips, 1963), was the first described and the most studied. Acetylation is mediated by lysine acetyltransferases (KATs) on lysine residues at N-terminal tail of core histones. There are two distinct families of HATs according to the property of their catalytic domains. One is the general control of amino acid synthesis protein 5-like 2 (GCN5)-related N-acetyltransferase (GNAT) family (Dyda et al., 2000). The other is the Moz-Ybf2/Sas3-Sas2-Tip60 (MYST) family, in which each member has a highly conserved MYST domain (Avvakumov and Cote, 2007; Pillus, 2008). Histone deacetylases (HDACs) are enzymes responsible for removing the acetyl group from histones (Brownell and Allis, 1996; Kuo and Allis, 1998; Roth et al., 2001). The orchestrated action of these two types of enzymes controls the reversible loop of this modification in organisms. Currently, there is a widely accepted neutralising theory for the mechanism of acetylation function. Acetylations on histones disrupt interactions between adjacent histones or contacts of histones and DNA by neutralising the basic charge of lysine residues, thus affecting chromatin to choose a more relaxed
conformation. Increasing evidence supports the theory that transcriptional outcomes are much more affected by the cumulative charge neutralization of histones by multiple lysine acetylations than by specific lysine acetylation (Zentner and Henikoff, 2013).

Hyperacetylated histones have been found to be associated with actively transcribed genes, facilitating transcription (Allfrey et al., 1964; Pogo et al., 1966). Researchers found that histone acetylation imparts in origin activation (Unnikrishnan et al., 2010), suggesting that acetylation is important for efficient DNA replication also by relaxing histone-DNA interactions. Furthermore, histone acetylation occurs at DNA double-strand breaks and may help increase the access of repair factors to DNA (Xu and Price, 2011).

Acetylated lysine residues are commonly recognised by the chromatin-associated proteins with bromodomain, such as, acetyltransferases and chromatin remodeling proteins (Zeng and Zhou, 2002).

Among the lysine acetylations, acetylation of histone H4 lysine 16 (H4K16ac) is of particular interest. H4K16ac is thought to be the foundation stone of histone H4 acetylation (Smith et al., 2003; Turner et al., 1992; Zhang et al., 2002). H4K16ac was initially discovered as an epigenetic mark in dosage compensation and catalysed by the histone acetyltransferase males absent on the first (MOF). In physiological conditions, the H4K16ac has been correlated with the activation of gene transcription, which loosens the compaction of 30 nm chromatin fiber to generate an open conformation of chromatin (Robinson et al., 2008; Shogren-Knaak et al., 2006). By DNA microarray analysis, specific transcriptional outcomes are detected, which are only caused by H4K16 mutation, totally independent of the mutation of the other lysine residues on the H4 tail (Dion et al., 2005). Only H4K16ac itself can reverse the silencing of mating-type genes in yeast (Johnson et al., 1990; Megee et al., 1990). The status of the H4K16 acetylation or deacetylation state can determine the boundaries of silenced chromatin (Kimura et al., 2002; Suka et al., 2002).

Both loss and increased levels of H4K16ac have been observed in connection
with cancer (Fraga et al., 2005; Kapoor-Vazirani et al., 2008; Pfister et al., 2008; Zhao et al., 2013). The forced reduction in of H4K16ac impairs DNA repair (Gupta et al., 2005; Li et al., 2010; Sharma et al., 2010; Taipale et al., 2005) and is connected to premature senescence in mice (Krishnan et al., 2011). Loss of H4K16ac in mouse T-cells results in their failure to differentiate and survive (Gupta et al., 2013). Moreover, H4K16ac is essential for embryogenesis (Gupta et al., 2008), and participates in the maintenance of mouse embryonic stem cell self-renewal and pluripotency (Li et al., 2012). Also in mice, H4K16 acetylation is found to be a key factor for purkinje cells’ survival and biological function, the loss of which leads to the neurological defects shown in ataxia telangiectasia (AT) patients (Kumar et al., 2011).

1.1.2.2 Histone phosphorylation
All core histones have been shown to be phosphorylated at their N-terminal domains. Histone phosphorylation plays important roles in numerous cellular processes, such as transcriptional regulation, cell cycle control, DNA repair and chromosome condensation (Banerjee and Chakravarti, 2011). Here, we will focus on the importance of the phosphorylation of the histone variant H2AX.

H2AX phosphorylation has been widely studied in the DNA damage response (DDR). The ataxia telangiectasia mutated (ATM), ataxia telangiectasia and Rad3 related (ATR) and DNA-dependent protein kinase (DNA-PK) kinases are known to phosphorylate H2AX at Ser139 in mammalian cells (Burma et al., 2001; Lowndes and Toh, 2005; Stiff et al., 2004; Ward and Chen, 2001). The H2AX phosphorylated at Ser139 is also called γH2AX. H2AX phosphorylation spans up to two megabases of DNA surrounding a double-strand break site and recruits DNA repair machinery to double-strand breaks to repair DNA damage. When DNA damage is repaired, H2AX is dephosphorylated by Wip1, a chromatin-associated phosphatase, allowing cells to re-enter the cell cycle (Macurek et al., 2010).

Interestingly, Kotova et al. reported that phosphorylation of H2Av, the
Drosophila homolog of the mammalian H2AX, recruits poly(ADP-ribose) polymerase 1 (PARP1) to chromatin, thus regulating expression of downstream genes (Kotova E, 2011).

1.1.2.3 Histone ubiquitination

Ubiquitination has been discovered in all four core histones, in which, histone H2A and H2B are most abundantly ubiquitinated. However, the precise functions of these modifications have not been elucidated yet.

In vertebrate cells, 5-15% of H2A and 1-2% of H2B are estimated to be modified by ubiquitin (Cao and Yan, 2012). Both H2A and H2B has been found to be targeted by mono- or poly-ubiquitination.

Monoubiquitination of lysine 119 on H2A (H2AK119ub) or lysine 120 on H2B (H2BK120ub) are the most studied. Both of them have significantly reduced levels in prostate and breast tumors, suggesting that they have some potential role in cancer (Prenzel et al., 2011; Zhu et al., 2007). Also they are involved in regulation of gene transcription. H2AK119ub, mediated by the B lymphoma Mo-MLV insertion region 1 (Bmi) protein, is frequently involved in gene silencing (Wang et al., 2004). H2BK120ub, modified by human Ring finger protein 20/40 (RNF20/RNF40) complex, is mostly correlated with active gene expression (Minsky et al., 2008; Pavri et al., 2006; Shema et al., 2008). Histone ubiquitination often crosstalks with other histone modifications to regulate transcription. H2BK120ub is required for both H3K4 methylation and H3K79 methylation by recruitment of SET1 and DOTL1 (Kim et al., 2009; McGinty et al., 2008; Wu et al., 2011). Recently, another H2B monoubiquitination, H2BK34ub, has been found in the trans-tail regulation of H3K4 and H3K79 methylation. This modification is mediated by human male specific lethal 2 (hMSL2) and human Male specific lethal 1 (hMSL1) proteins, the core of the human MSL complex (Hallacli et al., 2012). Crystal structure of nucleosomes suggests that this modification may promote Dot1L access to H3K79 via local nucleosome structure change. However, H2BK34 is not close to H3K4, so how
H2BK34 affecting H3K4 methylation is still not clear (Kim et al., 2009; Luger K., 1997; McGinty et al., 2008; Wu et al., 2011). Intriguingly, the ubiquitination of H2A and H2B by Ring finger protein 8 (RNF8) appears to promote H4K16ac in mouse spermiogenesis (Lu et al., 2010). H2A monoubiquitination participates in the random X chromosome inactivation in female mammals (de Napoles et al., 2004; Fang et al., 2004). Similar to non-histone proteins, polyubiquitination on different lysine residues on H2A and H2B can provoke different cellular events. For example, polyubiquitination on lysine 48 targets for proteasome mediated degradation (Chen et al., 1998; Liu et al., 2005), while K63-linked polyubiquitination of H2A and H2AX is usually in response to DNA damage repair (Bin and Elledge, 2007; Doil et al., 2009; Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007; Stewart et al., 2009).

1.2 Ubiquitination and RING finger domain E3 ligases

1.2.1 Ubiquitination

Ubiquitination is an enzymatic cascade process covalently transferring ubiquitin to target proteins, in which, three different types of enzymes; ubiquitin-activating (E1), ubiquitin-conjugating (E2), and ubiquitin-ligating (E3) participate (Komander, 2009).

Ubiquitin (Ub) is an 8 kD regulatory protein with 76 amino acids and has been found ubiquitously expressed in most tissues of eukaryotic organisms (Hershko and Ciechanover, 1998).

As shown in Figure 1.3, the process of ubiquitination comprises three main reactions (Lipkowitz and Weissman, 2011). First, Ub is activated in an ATP-dependent manner by E1 to generate a high-energy thiolester linkage between its C-terminal glycine residue and E1’s active cysteine site. Second, activated Ub on the E1 activating enzyme is transferred to the active-site cysteine of an E2 conjugating enzyme. Third, an E3 ligase binds to both the E2-Ub thioester and a protein substrate and brings them close to realise the final Ub transfer from E2-Ub to the protein substrate. Most frequently transferred Ub
links with the protein substrate via forming either an isopeptide bond between a protein lysine residue and the C-terminal glycine of ubiquitin or a peptide bond with the N-terminus of the substrate (Lipkowitz and Weissman, 2011). Furthermore, ubiquitination was found to occur on some other nucleophilic amino acids; threonine, serine and cysteine (Cadwell and Coscoy, 2005; Shimizu et al., 2010; Tokarev et al., 2011; Wang et al., 2007).

Figure 1.3 Ubiquitination cascade. A) Schematic representation of the three-step ubiquitination cascade. Monoubiquitination and K63 polyubiquitination generally serve non-proteolytic functions, whereas polyubiquitin chains of other linkages target proteins for degradation by the proteasome. B) Structure of ubiquitin highlighting its seven lysine residues (Liu and Chen, 2011).

In mammals, there are only two E1 enzymes, but ~50-60 different E2s and more than 600 E3s (Schulman and Harper, 2009). The E3 ubiquitin ligases are characterised into two types according to their catalytic domains, the RING domain E3 ligases (RING E3s) and the HECT domain E3 ligases (HECT E3s). The RING domain E3 ligases can be further subdivided into RING domain-containing E3 ligases and RING domain-related E3 ligases. The RING-related E3s include plant homeodomain (PHD) and leukemia associated protein (LAP) finger proteins and U-box family members (Deshaies and
Joazeiro, 2009). When these two types of E3s mediate Ub transfer, the HECT E3s need to form an obligate E3-Ub thioester intermediate, while the RING E3s, acting as a scaffold, directly transfer Ub from E2 to the substrate (Metzger et al., 2012).

The E3 ligases are vital for substrate specificity. A substrate can be monoubiquitinated by addition of an ubiquitin molecule or polyubiquitinated by addition of polyubiquitin chains to a lysine. It is generally considered that at least four lysine residues (11, 29, 48 and 63) on Ub act as linkage sites to form poly-Ub chains (Pickart, 2001; Wilkinson, 2000). In addition, multiple different lysine sites on a substrate can be monoubiquitinated in parallel, called multi-monoubiquitylation. Ubiquitination on different lysines results in different outcomes for the protein, such as proteasome-directed degradation, internalization and lysosomal targeting, alteration of subcellular localisation, promotion or disruption of protein-protein interaction, which in turn mediate cellular processes such as; regulation of transcription, DNA repair and transmembrane signalling (Acconcia et al., 2009; Bernasconi and Molinari, 2011; Metzger et al., 2012; Schile et al., 2008). Polyubiquitination on lysine 11, 29 and 48 have been found to target substrates for 26S proteasomal degradation. Lysine 48 polyubiquitination was found the most abundant in mammalian cells (Kim et al., 2011). However, monoubiquitylation or polyubiquitination on lysine 63 is usually associated with non-proteolytic events in intracellular trafficking, DNA repair, and signal transduction pathways (Grabbe and Dikic, 2009; Hicke et al., 2005).

Similar to other histone modifications, histone ubiquitination is a reversible process. deubiquitinases (DUBs) are required for removal of ubiquitin moiety. They are a class of thiol proteases, cleaving the isopeptide bond at glycine 76 on ubiquitin (Amerik and Hochstrasser, 2004; Atanassov et al., 2011; Reyes-Turcu et al., 2009).
1.2.2 RING domain E3 ligases

RING domain E3 ligases (RING E3s) are the most abundant E3s, and are conserved from yeast to human. There are at least 616 different RING E3s found in human cells, regulating many cellular processes and linked to multiple human diseases (Deshaies and Joazeiro, 2009). Despite their critical importance in biological function, their partners, substrates, and mechanism of action are not always clear.

The RING domain was first described by Freemont and colleagues, of which, the typical sequence is Cys-X_2-Cys-X_9-39-Cys-X_1-3-His-X_2-3-Cys/His-X_2-Cys-X_4-48-Cys-X_2-Cys (Freemont et al., 1991). Here, X represents any amino acid and the fifth histidine residue can be replaced by cysteine. Based on this, the RING motifs have been classified into two groups: C3HC4 RING (HC-type RING) and C3H2C3 RING (H2-type RING). The RING domain, via interaction with two zinc ions, forms a ‘cross-brace’ structure (Metzger et al., 2012).

RING E3s can form as monomers, dimers or integrate into multi-subunit complexes. To form homodimers or heterodimers, two molecules of RING E3s interact with each other through the RING finger domain or surrounding regions. In heterodimers, usually, only one RING domain exerts the ligase activity (Metzger et al., 2012). For example, casitas B-lineage lymphoma (c-Cbl) E3 ligase forms a homodimer and ubiquitinates activated receptor protein tyrosine kinases (RTKs) for degradation (Lipkowitz and Weissman, 2011). Murine double minute 2 (MDM2) E3 ligase can function as a homodimer or a heterodimer with MDMX, which mediates proteasomal degradation of p53 (Lee and Gu, 2010; Wade et al., 2010). The auto-ubiquitination of MDM2 itself was opted for regulation of its expression level (Fang et al., 2000). BRCA1, together with BRCA1-associated RING domain protein 1 (BARD1), mediates ubiquitination and is involved in the DNA damage response and DNA repair (Huen et al., 2010).
1.3 The MSL complex

1.3.1 The MSL complex in Drosophila

In various organisms, such as fruit flies, nematodes, and mammals, sex determination is controlled by heteromorphic sex chromosomes. To compensate for the gene dosage disequilibrium between males and females, diverse strategies have been developed in these organisms during evolution (Straub and Becker, 2007).

In Drosophila, dosage compensation is achieved by upregulating the X-linked gene expression 2-fold in males to ensure the equilibration in both sexes. The genes involved in this process were discovered by the genetic screen for mutants causing male-specific lethality. These genes include male specific lethal 1, 2 and 3 (MSL1, MSL2, MSL3), maleless (MLE) and males absent on the first (MOF) (Belote and Lucchesi, 1980; Hilfiker et al., 1997). Their protein products localize in the nucleus and, together with two noncoding RNAs (ncRNAs); roX1 and roX2, form the male specific lethal (MSL) complex (Figure 1.4), also known as the dosage compensation complex (DCC) (Gelbart et al., 2009; Rea et al., 2007).

Figure 1.4 The Drosophila MSL complex. The complex is composed of five proteins (MOF, MSL1, 2 and 3, and MLE) and two X-linked noncoding RNAs (roX1 and roX2) (Hosey and Brand, 2009).
Actually, most MSL proteins express in both sexes of flies, except for MSL2. Its expression is achieved only in male flies and has a pivotal role for the formation of the MSL complex (Kelley et al., 1995). MSL2 contains two conserved functional domains; an N-terminal RING finger domain and a C-terminal CXC (cysteine-rich) domain (Copps et al., 1998; Marin, 2003; Zhou et al., 1995). First, MSL2 interacts with MSL1 via its RING domain to form the core of complex, which binds to approximately 35-60 high-affinity sites along the length of the X chromosome (Dahlsveen et al., 2006; Gilfillan et al., 2007). The binding of MSL1/MSL2 promotes roX RNAs expression and initiates the MSL complex assembly. Finally, the complete MSL complex spreads to additional lower-affinity sites and coats the entire male X chromosome. (Meller and Rattner, 2002; Meller et al., 1997).

When the MSL complex is recruited to the male X chromosome, the hyperacetylation of H4K16 is enriched in the coding regions of transcriptionally active X-linked genes, which is mainly mediated by the MOF histone acetyltransferase. Interestingly, MSL3 is shown to directly bind to H3K36me3 and this histone marker is thought to be important for transcription elongation.

In females, to ensure the normal expression level of X-linked genes for viability, dosage compensation evolved to be silenced by inhibiting MSL2 expression. In this process, one key factor is female-specific RNA binding protein Sex-lethal (SXL). SXL mediates repression of MSL2 translation in both nucleus and cytoplasm. First, in the nucleus, SXL prevents MSL2 5’ untranslated region (UTR) splicing by preventing the U1/U2 small nuclear ribonucleoproteins (snRNPs) to bind to the splice sites, which in turn keeps a facultative intron in MSL2 5’ UTR. Then, in the cytoplasm, SXL binds to both the retained intron of 5’ UTR and the 3’ UTR for tightly translational repression of MSL2 translation (Graindorge et al., 2011).

The MSL complex is reported to interact with two nuclear pore components, Mtor/TPR and Nucleoporin 153 (Nup153) (Mendjan et al., 2006). Depletion of either of them causes loss of MSL’s X chromosome localization and dosage
compensation in male *Drosophila* (Mendjan et al., 2006). The Akhtar group further revealed that MOF is not only found on the X chromosome for the onset of dosage compensation but also on autosomes in a MSL complex-independent manner, regulating gene expression in *Drosophila* (Kind et al., 2008). This suggests that MOF’s function is probably beyond dosage compensation in *Drosophila*.

It was recently shown that *Drosophila* MSL2 could ubiquitylate MSL1, as well as MSL3 and MOF (Hallacli et al., 2012; Villa et al., 2012). This ubiquitylation can target these proteins for proteasome-mediated degradation to control MSL complex stoichiometry, but is also proposed to regulate their recruitment to specific chromatin domains (Villa et al., 2012).

### 1.3.2 The MSL complex in Human

Human orthologs of MSL proteins were found and they form an evolutionary conserved human MSL (hMSL) complex, also known as the hMOF (Males absent on the first, human) complex (Li et al., 2009; Mendjan et al., 2006; Smith et al., 2005). It is well known that humans have different dosage compensation mechanism compared to *Drosophila* and this complex does not appear to be involved.

The hMSL complex is composed of 5 proteins (hMOF, hMSL1, hMSL2, hMSL3, hMLE) and exhibits the same specificity for H4K16ac (Li et al., 2009; Smith et al., 2005).

#### 1.3.2.1 Human MOF

Human MOF (hMOF/MYST1/KAT8) is the human ortholog of the Drosophila MOF protein, which is ubiquitously expressed and is clearly targeted to all chromosomes (Smith et al., 2005). As one of the MYST (Moz-Ybf2/Sas3-Sas2-Tip60) histone acetyltransferase family members, hMOF contains a classic HAT dependent-MYST domain, but also possesses a chromodomain. Analysis of the subcellular localization of hMOF has shown that it localises to the nucleus (Neal
et al., 2000).

In 2005, the Roeder group reported that hMOF is stably associated with the histone methyltransferase MLL1 (mixed lineage in leukemia 1) in an independent complex. hMOF, together with MLL1, activates the Homeobox A9 (HOXA9) gene by facilitating both H4K16ac and H3K4me3 at the promoter. (Dou et al., 2005). One year later, Akhtar and co-workers also found that MOF was a key component in a discrete HAT complex named as the non-specific lethal (NSL) complex (Mendjan et al., 2006). In 2010, the findings from the Conaway group agreed that MOF is at least involved in two different multiprotein complexes, the MSL complex and the NSL complex (Cai et al., 2010).

The substrate specificity of hMOF is decided by the complex it is assembled into. Together with MSL members, hMOF specially acetylates H4K16 on nucleosomes. When associated with NSL members, hMOF not only exhibits the HAT activity on H4K16, but also can acetylate H4K5 and H4K8 with less specificity. In addition to histones, hMOF can take non-histone proteins as substrates. The lysine 120 of p53 protein is acetylated by the hMOF-associated NSL complex, which increases rapidly after DNA damage and activates proapoptotic gene expression (Li et al., 2009; Sykes et al., 2006). The TIP5 protein, the largest subunit of the chromatin remodelling complex NoRC, is also reported to be acetylated by hMOF on lysine 633 (Zhou et al., 2009). The deacetylase sirtuin-1 (SIRT1) leads to TIP5K633 deacetylation, enhancing promoter-associated RNA (pRNA) binding and an increase in heterochromatic histone marks. Reversible acetylation of the chromatin remodelling complex NoRC is required for non-coding RNA-dependent silencing. (Zhou et al., 2009). Recently, hMOF was reported to acetylate itself in the MYST domain. The lysine residue 274 is the major autoacetylated site (Lu et al., 2011; Yang et al., 2012).

Recent studies have also shown that hMOF plays a potential role in carcinogenesis. Pfister et al. found that hMOF and H4K16ac are frequently
down-regulated in primary breast carcinomas and medulloblastomas (Pfister et al., 2008). Consistent with this, the reduction or loss of H4K16ac in the majority of breast cancer tissues was reported by tissue microarray analysis (Elsheikh et al., 2009). Furthermore, reduced hMOF expression was detected in primary renal cell carcinoma (RCC), correlated with H4K16ac downregulation (Wang et al., 2013). hMOF was also found to contribute to non-small cell lung cancer (NSCLC) tumorigenesis. hMOF and H4K16ac are up-regulated in NSCLC tissue. And hMOF promotes NSCLC cell proliferation, migration and adhesion by mediating H4K16ac at the promoters of downstream target genes (Zhao et al., 2013). These findings suggest that hMOF expression in cancer can be up or down regulated, while the mechanism behind is not clear yet. hMOF was also reported to be involved in the autophagy pathway to determine cell survival or death. The reduction of hMOF and H4K16ac is associated with downregulation of autophagy-related genes (Fullgrabe et al., 2013).

Among the biological functions of hMOF, the role of hMOF in DNA damage repair is well studied. Taipale and coworkers found that hMOF knockdown cells display a significant delay in kinetics of DNA repair and are possibly involved in the ATM pathway (Taipale et al., 2005). The work of Pandita’s group identified the direct interaction between hMOF and ATM and suggests that functional hMOF participated in the autophosphorylation of ATM following Ionizing Radiation (IR), and further influenced the downstream factors of ATM (Gupta et al., 2005). hMOF mediated H4K16ac is required for recruitment of Mediator of DNA damage checkpoint 1 (MDC1) (Kruse and Gu, 2009). MOF depletion decreased DNA double-strand break repair by both NHEJ and HR by affecting IR-induced foci formation of γ-H2AX, Rad51, MDC1, 53BP1 and hSSB (Sharma et al., 2010). Consistent with this, overexpression of mouse Mof was found to improve the recruitment of 53BP1 to DNA damage sites in Zmpste24−/− MEFs (Krishnan et al., 2011). Recently, deacetylase SIRT1 was found to cooperate with hMOF in DNA damage response. When DNA damage was induced by IR or etoposide, the interaction between SIRT1 and hMOF
disrupted immediately, and reoccurred only six hours after DNA damage (Peng et al., 2012).

1.3.2.2 Human MSL1

hMSL1, like its *Drosophila* ortholog, has a conserved N-terminal coiled-coil domain and a C-terminal PEHE domain (Marin, 2003). The coiled-coil of hMSL1 interacts with the RING finger of hMSL2 (Li et al., 2005). The PEHE domain mediates the interaction with hMOF (Morales et al., 2004). Following the PEHE domain, there is a MSL3-interacting domain (Smith et al., 2005), that mediates binding with hMSL3.

Compared to the other two MSL members with enzymatic capabilities, hMSL1 is poorly understood.

Damokosh AI *et al.* developed a top 14-gene model to predict early recurrence of epithelial ovarian cancer after combination of platinum-paclitaxel chemotherapy. *hMSL1* gene was a good candidate and its upregulation was shown to be associated with earlier relapse (Hartmann et al., 2005). Furthermore, a gene-level screen was completed in 749 invasive ovarian cancer patients and 1041 controls. Results showed that two inherited single nucleotide polymorphisms (SNPs) within *hMSL1* were associated with risk of serous invasive ovarian cancer, in which, the SNP rs7211440 was correlated with decreased risk (Peedicayil et al., 2010). These observations implicate the regulation of hMSL1 in primary ovarian carcinoma.

The hMSL1 protein, similar to hMOF, was found to regulate global H4K16ac (Smith et al., 2005). hMSL1 was also found to interact with 53BP1, suggesting a potential role in DNA damage response (Gironella et al., 2009). To support this, mouse Msl1 was found important for MDC1 foci formation following IR (Li et al., 2010)

1.3.2.3 Human MSL2

Human male-specific lethal 2 (hMSL2) is composed of 577 amino acids,
contains a conserved N-terminal RING finger domain and C-terminal CXC domain (Marin, 2003; Smith et al., 2005). This RING finger protein is reported to be an E3 ubiquitin ligase.

Direct studies on hMSL2 are few; one study found that when overexpressed, it is able to monoubiquitinate p53 at lysine 351. This monoubiquitination targets the p53 protein for translocation to the cytoplasm, where it induces mitochondrial-dependent apoptosis (Kruse and Gu, 2009; Muscolini et al., 2011). Mutation of this K351 residue has been reported in a cisplatin-resistant ovarian carcinoma cell line (Muscolini et al., 2011). A second study has recently shown that hMSL2 in tandem with hMSL1 is able to ubiquitylate histone H2B on lysine 34 and this H2BK34ub directly regulates methylation of H3K4 and H3K79 by trans-tail crosstalk to promote transcription. Also, hMSL2 was found to be directly bound to *HOXA9* and *MEIS1* loci, activating their transcription through modification of local histones (Wu et al., 2011).

1.4 DNA damage

1.4.1 DNA damage response

Eukaryotic cells are continuously threatened with a large number of intrinsic and extrinsic DNA damage sources. Some intrinsic sources include reactive oxygen species (ROS), alkylating agents and replication collapse, while extrinsic sources include chemical agents, ultraviolet radiation (UV), ionizing radiation (IR). Failure to repair such lesions can lead to genomic instability and cancer (Ciccia and Elledge, 2010).

In order to keep genome integrity, eukaryotes have evolved a surveillance mechanism, termed as DNA damage response (DDR). DDR is a complex signal transduction process (Figure 1.5), including 3 key steps: DNA damage signal detection; DNA damage signals’ transduction; and the activation of cellular responses, such as cell-cycle arrests, DNA repair, apoptosis and senescence.
Figure 1.5 The DNA damage response pathway. The pathway is a complex signalling transducing cascade, including several levels of proteins, sensors, mediators and effectors (Sulli et al., 2012).

Sensor proteins directly recognise specific DNA lesions and activate the DDR (Zhou and Elledge, 2000). There are three main types of DNA damage sensors, the MRE11–RAD50–NBS1 (MRN) complex, replication protein A (RPA) and the RAD9–RAD1–HUS1 (9-1-1) complex. MRN complex detects DNA double-strand breaks (DSBs) (Lee and Paull, 2005), while RPA/9-1-1 complex recognizes the exposed regions of single-stranded DNA (ssDNA).

The sensed damage signalling is mainly transduced by the phosphatidylinositol 3-kinase-like protein kinases (PIKKs): ataxia-telangiectasia mutated (ATM), ataxia-telangiectasia and Rad3-related (ATR) and DNA-dependent protein kinase catalytic subunit (DNA-PKcs). In addition, the poly(ADP-ribose) polymerases (PARP) PARP1 and PARP2 are also considered as DNA damage transducers (Ciccia and Elledge, 2010). ATM exists as inactive dimers without DNA damage. When it is activated by DNA damage, ATM is recruited by the MRN complex to DSBs where it becomes dissociated, autophosphorylated and
acetylated (Bakkenist and Kastan, 2003, Kozlov 2006). ATR, bound by ATRIP, is recruited via the RPA/9-1-1 complex and activated along RPA-coated ssDNA regions formed at stalled replication forks and DSBs (Cimprich and Cortez, 2008). These in turn induce γH2AX formation in the vicinity of DNA lesions (Burma et al., 2001; Ward and Chen, 2001). DNA-PKcs is also activated after DNA damage, but only regulating a smaller group of proteins involved in DSB end joining. PARP1 and PARP2 can be activated by both single strand breaks (SSBs) and DSBs, adding poly(ADP-ribose) (PAR) chains on proteins to recruit downstream DDR factors at lesions (Schreiber et al., 2006).

Phosphorylated ATM, ATR and H2AX further recruit a number of mediators to amplify the DNA damage signals. The main mediators include mediator of DNA damage checkpoint 1 (MDC1), p53-binding protein 1 (53BP1), breast cancer 1 early-onset (BRCA1), topoisomerase II-binding protein 1 (TopBP1), claspin and Pax transactivation domain-interacting protein (PTIP) (Li and Zou, 2005). Here, we only discuss the first four. MDC1 is bound to γH2AX via its BRCT domains and amplifies H2AX phosphorylation, possibly by enforcing further activation of ATM and/or preventing H2AX dephosphorylation (Stucki and Jackson, 2006). MDC1 also establishes a positive feedback loop to spread γH2AX by stabilising the retention of NBS1 at DSBs (Kinner et al., 2008; Lukas et al., 2004). It has been found that IR-associated MDC1 and H2AX function as a scaffold and attracts additional factors to form IR-induced foci (IRIF) to damage sites (Lukas et al., 2004). MDC1 phosphorylation promotes RNF8-mediated ubiquitination of H2AX and possibly other proteins, which in turn recruits 53BP1 and the BRCA1 “Acomplex” (Bin and Elledge, 2007; Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007). 53BP1 is recruited to IRIFs in an H2AX- and MDC1-dependent manner (Harper and Elledge, 2007). 53BP1 can directly interact with MDC1 at the breaks and MDC1 depletion causes reduction of 53BP1 foci (Eliezer et al., 2009). The directed recognition of 53BP1 and γH2AX is important for 53BP1 accumulation on damage sites (Celeste et al., 2003; Ward et al., 2003). 53BP1 also was found to be involved in ATM
autophosphorylation and affects Chk2 and p53 activation (Wilson and Stern, 2008). BRCA1 is phosphorylated at Serine 1423, 1524 and other sites, and gets recruited to sites of DNA damage by ATM and ATR (Cortez et al., 1999; Tibbetts et al., 2000). BRCA1 phosphorylation is mainly ATM-dependent in response to IR, while phosphorylation in response to UV and hydroxyurea (HU) requires ATR (Tibbetts et al., 2000). TopBP1 binds the 9-1-1 complex and ATRIP, then is phosphorylated by ATR and stimulates ATR kinase activity (Kumagai et al., 2006). The activated mediators will further transfer the DNA damage signals to downstream effectors via the CHK1 and CHK2 kinases as well as other kinases such as CK2, p38, and MK2 (Harper and Elledge, 2007). CHK1 is mainly phosphorylated by ATR and CHK2 is mainly phosphorylated by ATM. Also some effectors can be directly phosphorylated by ATM/ATR in response to DNA damage. These effectors are composed of a large number of proteins, including the substrates of both PIK and CHK kinases and other proteins in DNA repair, transcription regulation and cell-cycle control, such as BRCA1, p53, Cdc25C and Nbs1(Zhou and Elledge, 2000).

1.4.2 DNA damage repair

As mentioned above, one cellular outcome following DNA damage signaling is DNA repair. The DNA damage effectors have evolved many elaborate pathways to repair different types of DNA lesions. Mismatched DNA bases are fixed by mismatch repair (MMR) (Jiricny, 2006). Small chemical alterations of DNA bases are corrected by base excision repair (BER) (Lindahl and Barnes, 2000). Pyrimidine dimers and intrastrand crosslinks are corrected by nucleotide excision repair (NER) (Ciccia and Elledge, 2010). Single-strand breaks (SSBs) are processed by single-strand break repair (SSBR) (Caldecott, 2008). DNA double-stranded breaks (DSBs) are repaired by homologous recombination (HR) and nonhomologous end joining (NHEJ) (West, 2003). Here, we will focus on DNA double-stranded break induced repair pathways.
DSBs are some of the most harmful lesions, caused by endogenous metabolic products and reactive oxygen species, and also exogenous IR (Pardo et al., 2009). They also occur when dicentric chromosomes are pulled to opposite spindle poles during mitosis (Pardo et al., 2009). DSBs are more dangerous to genomic integrity and more difficult to repair because both strands of DNA are broken, offering no intact template strand for repair synthesis (Rothkamm et al., 2003). Incorrect repair of DSBs are lethal as they result in chromosomal breakage, fragmentation and translocation. Eukaryotic cells have evolved two distinct but interconnected systems for DSBs repair: homologous recombination (HR) and non-homologous end-joining (NHEJ) (Figure 1.6). NHEJ promotes the potentially inaccurate religation of DSBs. HR precisely restores the genomic sequence of the broken DNA ends.

DSBs are repaired mainly via HR in bacteria and yeast (Dudasova et al., 2004; Kanaar et al., 2008b). In mammalian cells, more than 90% of DSBs are repaired by NHEJ throughout the cell cycle, whereas, HR is generally restricted to late S and G2 phases of the cell cycle (Kanaar et al., 2008b; Rothkamm et al., 2003). The choice of HR or NHEJ depends mainly on cell cycle stage (Rothkamm et al., 2003; Sonoda et al., 2006). The HR pathway requires homologous DNA duplex as a repair template. The homologous chromosomes/sister chromatids, which only appear during the S and G2 phases are the preferred HR template. Cohesins, a ring-shaped complex which maintain sister chromatids in tight proximity, are recruited to DSB site by γH2AX (Kim et al., 2002). Cohesins mediate close association of sister chromatids and facilitate homology searching and subsequent repair. Proteins that promote DSB-end resection, such as CtIP, are up-regulated in S and G2 phases (Pardo et al., 2009). The up-regulated resection in turn activates HR accessibility. However, NHEJ remains functional in S and G2 phases. NHEJ is an error-prone repair pathway which can cause the harmful replication fork stalling or collapse during DNA replication (Adachi et al., 2004). Hence, cells may somehow down-regulate NHEJ or suppress the toxic effects of NHEJ (Pardo et al., 2009).
In addition to repair of DSBs, the HR pathway is also involved in repair of DNA interstrand crosslinks and DNA gaps (Pardo et al., 2009), whereas NHEJ is also essential for the Ig V(D)J and class-switch recombination in the development of the T and B-cell repertoires (Hefferin and Tomkinson, 2005).

Figure 1.6 DNA repair pathways involved in the repair of double-strand breaks. A) Rapid association of Ku to DSBs promotes NHEJ by recruiting DNA-PKcs. B) In HR, MRN recruited to DSBs by PARP in competition with Ku, promotes homologous recombination in S and G2. C) Limited DSB resection carried out by CtIP and MRN in G1 results in alternative NHEJ (Ciccia and Elledge, 2010).
1.4.2.1 Homologous recombination repair pathway (HR)

HR reactions start with the 5’-3’ exonuclease-mediated degradation of 5’-ended DNA strand at DSBs. Following 5’-end resection, a 3’ single-stranded DSB end is generated. It can search and invade the homologous template and functions as a primer during repair DNA synthesis. The 3’ single-stranded DSB end invades and replaces one strand of homologous DNA duplex (strand exchange) and pairs with the other strand resulting in formation of a heteroduplex or hybrid DNA called displacement-loop (D-loop). From now on, several HR sub-pathways are employed to replace the fragment containing DSB with newly synthesized sequence (Hartlerode and Scully, 2009).

HR repair is a multistep process that involves numerous proteins and enzymes. There are some different sub-pathways depending on strand invasion and DNA synthesis primer abilities of both single-stranded DSB ends at the DSB site. However, most of the enzymatic steps are shared by the different sub-pathways (Li and Heyer, 2008). The MRN complex and nuclease EXO1 are responsible for the resection of DSBs to initiate HR repair (Pardo et al., 2009). RAD51 together with RPA promotes homology search and DNA strand exchange of the single-stranded DSB end (Sung, 1994). Loading of RAD51 onto damaged DNA is facilitated by RAD52 or/and BRCA2 (Esashi et al., 2007; Song and Sung, 2000). Rad52 also contains resolvase activity for HJ cleavages (Miyazaki et al., 2004). DNA helicases, topoisomerases, polymerases, nucleases and ligases catalyze the DNA synthesis to restore the lost genomic information at the DSB site and the completion of the repair process (Cejka et al., 2010; Filippo et al., 2008). Other members of the RAD52 epistasis group, such as RAD54, RAD55, RAD57, RAD59, RAD50 also participate in the HR repair (Symington, 2002).

1.4.2.2 Non-homologous end joining pathway (NHEJ)

In NHEJ pathway, almost any kind of DSB ends are directly rejoined together following little or no nucleolytic processing of the ends. As it does not require a homologous sequence, small insertions and deletions may occur after repair.
These mutations usually happen when the two DSBs are not fully compatible. Several (1-4) terminal bases may be degraded before the two ends are paired (Daley and Wilson, 2005). The repair is finished by gap filling synthesis and ligation (Pardo et al., 2009). If the two DSB ends at the DSB location are complementary or blunt from the same parental DNA duplex, the repair of the DNA is faithful.

Like HR repair, many protein factors are implicated in the NHEJ reactions. The MRX and DNA-PK complexes first bind DSB ends to maintain DSB ends close to each other and recruit the ligase complex (Hopfner et al., 2002). The DNA-PK complex comprises a catalytic subunit DNA-PKcs and a heterodimeric complex (Ku) of KU70 and KU80. Ku's function is to protect DSB ends from degradation and to recruit the ligase complex to the DSB site (Ribes-Zamora et al., 2007). DNA-PKcs is a serine/threonine kinase which autophosphorylates itself and several other DDR components (Collis et al., 2005). It also helps bridge both DNA ends together through formation of synaptic complex (Spagnolo et al., 2006). Before ligation commences, some mismatched and incompatible DSB ends such as ionizing radiation-induced DSBs, must be corrected and re-synthesized for single-stranded gap filling. The mammalian polynucleotide kinase (PNK) and Aprataxin coupled with XRCC4 and DNA nucleases perform the correction process (Chappell et al., 2002; Clements et al., 2004). The single-strand gaps outside the partial paired overhangs are filled by DNA polymerases. The unpaired flap DNA is cleaved by Artemis through interaction with DNA-PK (Niewolik et al., 2006). Finally, ligase IV in a complex with XRCC4 or XLF carries out the relegation of the two repaired ends (Ahnesorg et al., 2006; Grawunder et al., 1997).

1.4.2.3 Microhomology-mediated end joining (MMEJ or alt-NHEJ)

Recently, a new Ku-independent end joining pathway, termed microhomology-mediated end joining (MMEJ) or alternative NHEJ (alt-NHEJ), has been proposed, which functions as backup to the classical NHEJ pathway (Ma et al.,
When the classical NHEJ pathway is chemically or genetically impaired, cells do not shift DSBs to HR repair but instead use MMEJ repair. The mechanism of this pathway has not been well elucidated. MMEJ has shown higher activity in G2 phase that cannot be attributed to HR. It only requires imperfect microhomology (5-20 nucleotides in length), so compared to NHEJ, a larger sequence is lost after MMEJ repair, indicating an error-prone repair. DNA ligase III, PARP1 and histone H1 are found involved in this pathway (Iliakis, 2009). DNA ligase III forms a complex with XRCC1, which is regulated by PARP1 (McKinnon and Caldecott, 2007). This implies that a long 3’ single-stranded DSB end forms in MMEJ pathway. Thereafter, components of both NHEJ and HR are involved in MMEJ (Decottignies, 2007; Ma et al., 2003).

1.4.3 Histone modifications in DNA damage response and repair

Accumulating evidence has linked many types of histone modifications to the DNA damage signalling and repair process. These modifications not only transduce damage signals, but also recruit specific subsets of factors to respond to and repair DNA damage.

One of the best studied histone modification in the DNA damage pathway is γH2AX. It is well known that this early event is mediated by ATM, ATR and DNA-PK kinases (Rogakou et al., 1998). γH2AX has been shown to be not strictly necessary for the initial recruitment of the DDR machinery, but essential for the amplification of the DDR signal and the formation and maintenance of DDR foci (Celeste et al., 2003). When repair is completed, dephosphorylation of γH2AX is a prerequisite to turn off checkpoint signalling (Keogh et al., 2006). In mammalian cells, γH2AX levels reach a peak within 1-2 hours following damage, and steadily decreases over a time course of 12 hours in accordance with DNA repair kinetics (Paull et al., 2000). The protein phosphatase 2A (PP2A) and wild type p53-induced phosphatase (Wip1) are required for γH2AX dephosphorylation (Cha et al., 2010; Chowdhury et al., 2005; Macurek et al., 2003).
2010). Interestingly, concomitantly with an increase in γH2AX levels after DNA damage, another H2AX modification, phosphorylation of tyrosine 142 is lost, which is constitutively phosphorylated in normal conditions (Sulli et al., 2012). The mediator MDC1 can only bind dephosphorylated tyrosine 142 in H2AX (Wang et al., 2011).

Following DNA damage, H2AX is not only phosphorylated but also undergoes ubiquitination by ring finger protein 2 (RNF2), which recruits B lymphoma Mo-MLV insertion region 1 homolog (BMI1) to sites of DNA lesions (Sulli et al., 2012). The recruitment of BMI1 ensures the proper localization of 53BP1, BRCA1 and RAP80 to DNA DSBs (Ismail et al., 2010). BMI1 also blocks transcriptional elongation at the DNA damage sites (Chagraoui et al., 2011). BMI1 is also a component of polycomb repressive complex 1 (PRC1), which ubiquitylates histone lysine 119 on H2A a modification associated with DNA breaks (Ginjala et al., 2011; Ismail et al., 2010). In addition to RNF2, the E3 ligases RNF8 and RNF168 mediate the ubiquitination of γH2AX. First, recruitment of RNF8 to DSB sites promotes initial K63-linked ubiquitination. Then, RNF168 amplifies and spreads polyubiquitination to maintain the retention of repair factors at the damaged chromatin (Bekker-Jensen and Mailand, 2010; Panier and Durocher, 2009; Ulrich and Walden, 2010). It was found that the acetylation of Lysine 5 in H2AX by 60 kDa tat-interactive protein (TIP60) is necessary for H2AX ubiquitination (Ikura et al., 2007). The RNF8 and RNF168 also participate in the addition of ubiquitin to histones H2A and H2B, enhancing the localization of 53BP1 and BRCA1 at DSBs (Doil et al., 2009; Mailand et al., 2007; Stewart et al., 2009; Wu et al., 2009). In the RNF8 mediated ubiquitination process, an E3 ligase HERC2 is recruited to repair foci, somehow facilitating the interaction of RNF8 with the E2 enzyme Ubc13 (Bekker-Jensen et al., 2010; Plans et al., 2006).

Histone H2B is also subject to monoubiquitination on lysine 120 by the RNF20/RNF40 complex in response to DSB damage (Shiloh et al., 2011). Interference with H2B ubiquitylation perturbs XRCC4 and Ku80 recruitment,
finally impairing DSB end resection and Rad51 accumulation (Moyal et al., 2011; Nakamura et al., 2011). Additional histone phosphorylations in DSB repair have been described. Serine 14 of Histone H2B is phosphorylated in DNA damage and this is thought to be involved in chromatin condensation and apoptosis (Fernandez-Capetillo et al., 2004). When DSBs exist, the phosphorylation of serine 1 on histone H4 by casein kinase 2 is inhibitory to NuA4-mediated H4 acetylation (Cheung et al., 2005; Utley et al., 2005).

H3K9me3 is commonly considered as the marker of heterochromatin formation and gene silencing. Heterochromatin is often correlated with suppression of DDR signaling. Unexpectedly, this modification plays a role in the activation of ATM activity. In an H3K9me3-dependent manner, TIP60 is directed to DNA damage sites and activates ATM through acetylation (Ayoub et al., 2008; Sun et al., 2009).

Except for H3K9me3, two other histone methylations, H3K79me and H4K20me, are better studied in DNA damage response. H3K79 gets methylated by histone methyltransferase DOT1L, and gets bound by 53BP1 via its Tudor domain (Huyen et al., 2004). As the levels of methylated H3K79 keep static in response to DSB damage, it has been suggested that the exposure of methylated H3K79 is more likely serving as a sensor showing the abnormal chromatin relaxation state (Huyen et al., 2004).

H4K20me1, H4K20me2, and H4K20me3 are also increased in the vicinity of DNA lesions. The histone methyltransferase MMSET, responsible for dimethylation and trimethylation, has been found to be recruited to DNA damage sites in a γH2AX–MDC1-dependent manner, involving 53BP1 recruitment (Pei et al., 2011).

However, there is no clear evidence showing that any of these histone modifications are specific for a specific DNA repair pathway. Rather than their appearance, these histone modifications’ disappearance or maintenance seem to privilege a given repair pathway. The deacetylation of various histone residues seems to facilitate the religation step of NHEJ, as the recombination process
required the retention of an open (acetylated) chromatin structure. Deacetylation is also needed later during HR for restoration of the initial chromatin structure and to avoid unnecessary recombination events (Escargueil et al., 2008).

1.5 Chicken DT40 cell line as a model system
Gene targeting and RNAi methods are widely used to interpret gene function in vertebrate and mammal cells. RNAi knockdown is less time consuming compared to gene targeting, but how to avoid off-target effects with siRNAs remains a problem for researchers. RNAi does not abolish the transcript completely, and a gene’s function is probably underestimated if the gene is effective at low expression levels or its effects are somehow complemented by related genes (Kohonen et al., 2007). So, gene targeting based on homologous recombination is considered a powerful method to precisely manipulate the genome. It is more laborious but it does completely shut down gene expression. DT40 B lymphocyte cell line is one of the most valuable and widely used vertebrate cell lines for gene targeting.

1.5.1 Chicken DT40 cell line
The DT40 B lymphocyte cell line was originally derived from an avian leukosis virus-transformed bursa of Fabricius from a Hyline SC female chicken (Baba et al., 1985; Baba and Humphries, 1984). DT40 cells are small, approximately 10 mm in diameter with a high nucleus to cytoplasm ratio (Winding and Berchtold, 2001). In general, DT40 cells grow rapidly as suspension cells with a doubling time of 8-10 hours. DT40 cells have been reported to express high levels of the c-myc protein as they were generated with transfection of the avian leukosis virus, which keeps them undergoing immunoglobulin light chain gene conversion (Buerstedde et al., 1990; Thompson et al., 1987). DT40 cells display a relatively stable karyotype with 80 modal chromosomes, including 11 autosomal macrochromosomes, 67 microchromosomes and the heterogametic ZW sex chromosomes (Sonoda et al., 1998). The DNA content of
the largest macrochromosomes is 20 times more than that of the minichromosomes (Johnston et al., 2010). Spontaneous breaks are only found in one percentage of nuclei (Hochegger et al., 2004). Among all these chromosomes, most of the chromosomes are disomic, except for chromosome 2 and 4. Chromosome 2 and one additional microchromosome are trisomic and chromosome 4 is monosomic (Sonoda et al., 1998).

With the sequencing of most of the chicken genome in 2004 (Burt and Pourquie, 2003), we know that the chicken genome is only one-third the size of mammalian genomes (Hudson et al., 2002). The genes are normally smaller than their mammalian orthologs, which facilitates gene targeting. However, there is one thing we need to keep in mind; the current genome assembly still contains gaps affecting 5-10% of all transcripts.

As a cell-line knockout model, DT40 cells show a number of advantages. The most important is that DT40 cell line has a high frequency of homologous recombination to mediate efficient gene targeting. The frequency varies between 10 and 80%, several orders of magnitude higher than in mammalian cells (Buerstedde and Takeda, 1991). Also in contrast to mammalian systems, DT40 knockout takes less time to achieve and is less costly. Furthermore, the absence of functional p53 in DT40 cells facilitates studies of DNA damage due to the silenced damage-induced apoptosis during interphase (Yamazoe et al., 2004). So far, the DT40 cell line has been used to study B-cell receptor (BCR) signalling (Alinikula et al., 2006), transcriptional regulation of B-cell function (Arakawa and Buerstedde, 2004), DNA repair (Morrison and Takeda, 2000; Yamazoe et al., 2004), chromatin organization and recombination (Nakayama and Takami, 2001; Winding and Berchtold, 2001).

1.5.2 Gene targeting in DT40 cell line

For successful integrated targeting of a gene of interest, the standard strategy is to design and transfect optimal disruption constructs into DT40 cells to disrupt gene expression by homolog recombination events. In cells, the disruption
constructs function in two ways: one is to introduce the selection cassette into a specific site of the coding sequences of the target gene, and the other is to replace a partial or full gene coding region. A targeting construct consists of at least 2 kb of chicken homologous genomic DNA on either side of the selection cassette. It has been found that increasing the length of flanking DNA sequence exhibits higher targeting frequency. However, shorter homolog arms have also been successfully used (Sonoda et al., 1998; Takata et al., 1995). So far, there are six selection cassettes used, including puromycin, blasticidin S, hygromycin, histidinol, neomycin and gpt (Lahti, 1999; Yamazoe et al., 2004). They are modified to be compatible with the Cre/lox system. Hence, these resistant cassettes can be recycled, reducing any potential disruption to gene locus since they are excised from the genome (Arakawa et al., 2001). Also, the recycling of cassettes makes several rounds of gene disruption possible with only one drug. However, for the essential genes, the standard strategy cannot be applied. To keep the cell alive, a conditional knockout strategy should be used.

1.6 Aim of the thesis

The overall aim of this study is to better understand cellular function(s) of the poorly characterised Msl2 in the vertebrate system.

To achieve this, our first goal was to generate Msl2 knockout cells and the appropriate controls (see Chapter 3). Our second goal was to use a loss of function approach, subjecting these cells to a series of assays, to elucidate the function of Msl2.
Chapter 2 Materials and methods

2.1 Materials

2.1.1 Solutions, buffers and media

Most common chemicals used in this study were purchased from Sigma, Fisher Scientific or GE Healthcare unless specified below or in the text. All solutions, buffers and media frequently used in this thesis are listed in Table 2.1. They were prepared with double-distilled H2O (ddH2O) or MilliQ water (Millipore) and autoclaved or filtered prior to use if necessary.

Table 2.1 Common reagents and buffers

<table>
<thead>
<tr>
<th>Name</th>
<th>Composition</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>6× DNA loading dye</td>
<td>20% Sucrose, 0.1 M EDTA pH 8.0, 1% SDS, 0.25% Bromophenol blue, 0.25% Xylene cyanol</td>
<td>DNA sample loading</td>
</tr>
<tr>
<td>1× TAE</td>
<td>40 mM Tris-acetate pH 8.0, 1 mM EDTA</td>
<td>Agarose gel running</td>
</tr>
<tr>
<td>Tail buffer</td>
<td>50 mM Tris pH 8.8, 100 mM EDTA, 100 mM NaCl, 1% SDS</td>
<td></td>
</tr>
<tr>
<td>Maleic acid wash buffer</td>
<td>0.1 M Maleic acid pH 7.5, 0.15 M NaCl, 0.3% Tween 20</td>
<td></td>
</tr>
<tr>
<td>Blocking solution</td>
<td>Maleic acid wash buffer, 10% Caseine</td>
<td></td>
</tr>
<tr>
<td>Denaturation solution</td>
<td>1.5 M NaCl, 0.5 M NaOH</td>
<td></td>
</tr>
<tr>
<td>Depurination solution</td>
<td>250 mM HCl</td>
<td>Southern blot</td>
</tr>
<tr>
<td>Neutralisation solution</td>
<td>0.5 M Tris-HCl pH 7.5, 1.5 M NaCl</td>
<td></td>
</tr>
<tr>
<td>10× SSC</td>
<td>1.5 M NaCl, 0.15 M Na citrate, pH adjust to 7.0 with citric acid</td>
<td></td>
</tr>
<tr>
<td>High stringency buffer</td>
<td>0.5× SSC, 0.1% SDS</td>
<td></td>
</tr>
<tr>
<td>Low stringency buffer</td>
<td>2× SSC, 0.1% SDS</td>
<td></td>
</tr>
<tr>
<td>Detection buffer</td>
<td>0.1 M Tris pH 9.5, 0.1 M NaCl</td>
<td></td>
</tr>
<tr>
<td>4× Laemmli buffer</td>
<td>150mM Tris pH 6.8, 9% SDS, 0.03% bromophenol Blue, 10%β-mercaptoethanol</td>
<td>Protein sample loading</td>
</tr>
</tbody>
</table>
## Chapter 2 Materials and methods

<table>
<thead>
<tr>
<th>Buffer/Solution/Buffer <strong>Stacking gel buffer</strong></th>
<th>0.5M Tris pH 6.8</th>
<th>SDS-PAGE analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Separating gel buffer</strong></td>
<td>1.5M Tris pH 8.8</td>
<td></td>
</tr>
<tr>
<td><strong>1× Running buffer</strong></td>
<td>0.025 M Tris, 0.25 M glycine, 0.1% SDS</td>
<td></td>
</tr>
<tr>
<td><strong>Coomassie brilliant blue R</strong></td>
<td>0.5% Coomassie, 35% Methanol, 14% acetic acid</td>
<td></td>
</tr>
<tr>
<td><strong>Coomassie destain solution</strong></td>
<td>30% methanol, 10% acetic acid</td>
<td></td>
</tr>
<tr>
<td><strong>1× Transfer buffer</strong></td>
<td>48 mM Tris, 39 mM glycine, 20% methanol, 0.0375% SDS</td>
<td>Western blot</td>
</tr>
<tr>
<td><strong>Ponceau S solution</strong></td>
<td>0.5% Ponceau S, 5% acetic acid</td>
<td></td>
</tr>
<tr>
<td><strong>Blocking solution</strong></td>
<td>1× PBS, 0.1% Tween 20, 2-5% skimmed milk or 1% BSA</td>
<td></td>
</tr>
<tr>
<td><strong>PBST</strong></td>
<td>1× PBS, 0.1% Tween 20</td>
<td></td>
</tr>
<tr>
<td><strong>Antibody dilution buffer</strong></td>
<td>1× PBS, 0.1% Tween 20, 1% skimmed milk or 1% BSA</td>
<td></td>
</tr>
<tr>
<td><strong>1× Phosphate buffer saline (PBS)</strong></td>
<td>137 mM NaCl, 2.7 mM KCl, 1.4 mM NaH2PO4, 4.3 mM Na2HPO4, pH 7.4</td>
<td>Washing cells</td>
</tr>
<tr>
<td><strong>RIPA buffer</strong></td>
<td>50 mM Tris-HCl pH ?, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS</td>
<td>Whole cell protein extraction</td>
</tr>
<tr>
<td><strong>Isotonic lysis buffer</strong></td>
<td>10 mM Tris pH 7.5, 2 mM MgCl2, 3 mM CaCl2, 0.32 M sucrose supplemented with protease inhibitors and 1 mM DTT</td>
<td>Subcellular fraction extraction</td>
</tr>
<tr>
<td><strong>Extraction buffer</strong></td>
<td>20 mM HEPES pH 7.7, 1.5 mM MgCl2, 0.42 M NaCl, 0.2 mM EDTA, 25% (v/v) Glycerol supplemented with protease inhibitors and 1 mM DTT</td>
<td></td>
</tr>
<tr>
<td><strong>Blocking solution</strong></td>
<td>1× PBS, 1% BSA (sterile)</td>
<td>Immunofluorescence</td>
</tr>
<tr>
<td><strong>Fixation solution</strong></td>
<td>1× PBS, 4% Paraformaldehyde</td>
<td></td>
</tr>
<tr>
<td><strong>Permeabilisation buffer</strong></td>
<td>1× PBS, 0.15% Triton-X 100</td>
<td></td>
</tr>
<tr>
<td><strong>Luria-Bertani (LB) Broth</strong></td>
<td>1% tryptone, 0.5% yeast extract, 1% NaCl, pH adjusted to 7.0 with 4 M NaOH</td>
<td>E.coli growth</td>
</tr>
<tr>
<td><strong>semi-solid methylcellulose medium</strong></td>
<td>1.5% methylcellulose, 1× DMEM/F-12 with L-glutamine (+), 15% FBS, 15% chicken serum, 5% Penicillin/Streptomycin, 50 μM β-mercaptoethanol</td>
<td>Survival assay</td>
</tr>
</tbody>
</table>
2.1.2 Molecular biology reagents

All of the restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (NEB), Shrimp Alkaline Phosphatase (SAP) from Promega, DNA polymerases KOD Hot Star and Taq from Takara and Sigma. DNA and protein size markers were bought from Invitrogen and Bio-rad. Molecular biology kits bought from various companies are listed in Table 2.2.

Table 2.2 Molecular biology kits in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Use</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Pure PCR Cleanup Micro Kit</td>
<td>Purification of DNA fragments</td>
<td>Roche</td>
</tr>
<tr>
<td>QIAquick Gel Extraction Kit</td>
<td>Extraction of DNA bands from agarose gel</td>
<td>Qiagen</td>
</tr>
<tr>
<td>NucleoBond Xtra Midi Plus</td>
<td>Large scale extraction of plasmid DNA</td>
<td>Macherey-Nagel</td>
</tr>
<tr>
<td>ISOLATE RNA Mini kit</td>
<td>RNA extraction</td>
<td>Bioline</td>
</tr>
<tr>
<td>High Capacity RNA-to-cDNA Kit</td>
<td>cDNA synthesis</td>
<td>Applied Biosystems</td>
</tr>
<tr>
<td>Fast SYBR Green Master Mix</td>
<td>Real-time PCR</td>
<td>Applied Biosystems</td>
</tr>
<tr>
<td>PCR DIG Probe Synthesis Kit</td>
<td>Southern blot probe labelling</td>
<td>Roche</td>
</tr>
<tr>
<td>pGEM-T Easy Vector kit</td>
<td>Cloning</td>
<td>Promega</td>
</tr>
</tbody>
</table>

Primers for cloning and realtime PCR used in this study are listed in Table 2.3. All of them were synthesised in Sigma.

Table 2.3 Primers used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Msl2 targeting 5’ Arm</td>
<td>GATCTGGTACTTTGAGAGCCTGTG CAGATGTGAGTGAACTGCAAGAGAT</td>
</tr>
<tr>
<td></td>
<td>actagTTTGAGATGAAATTGCTGATGTAAATG</td>
</tr>
<tr>
<td></td>
<td>acgcgtCAAATGCTGAAGTAGAACTGCTGCA</td>
</tr>
<tr>
<td>Msl2 targeting 3’ Arm</td>
<td>GGAATGGTGTTAGAAGTATTACAG CTAACCACATCTCTAAACCAAAG</td>
</tr>
<tr>
<td></td>
<td>ctcgagctATGAACCCGCTGAATGCCA GCT</td>
</tr>
<tr>
<td>Msl2 targeting 5’ probe</td>
<td>gaattcTCAACAGTCATATCTCACGTCCTATA</td>
</tr>
<tr>
<td>Msl2 cDNA cloning</td>
<td></td>
</tr>
<tr>
<td>Gene (qPCR)</td>
<td>Forward Primer</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Msl2 (C44R)</td>
<td>TCCCTCTCCCGCTGCCTGTG</td>
</tr>
<tr>
<td>Msl2 (ΔRING)</td>
<td>aaaaagcttGAGTACATAACAA</td>
</tr>
<tr>
<td>Msl2 qPCR</td>
<td>TCTAATTTGCAAGCCAGCTT</td>
</tr>
<tr>
<td>Mof qPCR</td>
<td>GGAAATCTACCGCAAGAGCAAC</td>
</tr>
<tr>
<td>Dot1l qPCR</td>
<td>GACAAAACACCATGATCGCTGCTATG</td>
</tr>
<tr>
<td>Suv420 qPCR</td>
<td>TGGCTTGTCTGTTCGCTG</td>
</tr>
<tr>
<td>Hoxa9 qPCR-1</td>
<td>GCAAAACGGAGAAACCACAC</td>
</tr>
<tr>
<td>Hoxa9 qPCR-2</td>
<td>ACGGCAGGTACATCGTCTCGT</td>
</tr>
<tr>
<td>Meis1 qPCR</td>
<td>TGTAGCTTTCCCAGCAGCAG</td>
</tr>
<tr>
<td>Gapdh qPCR</td>
<td>GCAGATGCAGGTGCTGAG</td>
</tr>
<tr>
<td>β-Actin qPCR</td>
<td>TCATCACCATTGGCAATGAGGA</td>
</tr>
<tr>
<td>hMSL1 qPCR</td>
<td>CAGGCCAAAGAAAGGAGAT</td>
</tr>
<tr>
<td>MDC1 qPCR</td>
<td>GAGGGGAAGTCTCCAGAAG</td>
</tr>
<tr>
<td>hMSL3 qPCR</td>
<td>GTGCTTCTGTGATCCGATGA</td>
</tr>
<tr>
<td>GAPDH qPCR</td>
<td>ATCCACCCATGGCAATATC</td>
</tr>
<tr>
<td>PPIA qPCR</td>
<td>AGGGTTTCTGCCTTCTACAGA</td>
</tr>
<tr>
<td>MOF qPCR</td>
<td>AGATCTACCGCAAGAGCAACA</td>
</tr>
<tr>
<td>MSL2 qPCR</td>
<td>GTCCTTTTCGTGCTGTTT</td>
</tr>
</tbody>
</table>
Vectors and plasmids used in this study for cloning, gene targeting or protein expression in prokaryotic or eukaryotic cells were listed in Table 2.4.

**Table 2.4 List of vectors and plasmids**

<table>
<thead>
<tr>
<th>Name</th>
<th>Use</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGEMT-Easy vector</td>
<td>Cloning</td>
<td>Promega</td>
</tr>
<tr>
<td>pGEMT-Msl2-Puro</td>
<td>Gene targeting</td>
<td>This study</td>
</tr>
<tr>
<td>pGEMT-Msl2-Blast</td>
<td>Gene targeting</td>
<td>This study</td>
</tr>
<tr>
<td>pCDNA3.1-HA-2F vector</td>
<td>Cloning</td>
<td>Simona Moravcová</td>
</tr>
<tr>
<td>pCDNA3.1-HA-2F-Msl2 (rescue)</td>
<td>Generation of stable cell line</td>
<td>This study</td>
</tr>
<tr>
<td>pCDNA3.1-HA-2F-Msl2 (C44R)</td>
<td>Generation of stable cell line</td>
<td>This study</td>
</tr>
<tr>
<td>pCDNA3.1-HA-2F-Msl2 (ΔRING)</td>
<td>Generation of stable cell line</td>
<td>This study</td>
</tr>
<tr>
<td>pGEX-6P2-GST-RNF168</td>
<td>Bacterial expression</td>
<td>Dr. Lorenza Penengo</td>
</tr>
<tr>
<td>pET41b-GST vector</td>
<td>Bacterial expression</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pET41b-GST-hMOF</td>
<td>Bacterial expression</td>
<td>Dr. Mikko Taipale</td>
</tr>
</tbody>
</table>

**2.1.3 Bacterial strains and cell lines**

**2.1.3.1 Bacterial strains**

The bacterial strains used in this study are listed in Table 2.5.

**Table 2.5 Strains of E. coli used in this study**

<table>
<thead>
<tr>
<th>E. coli strain</th>
<th>Genotype</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top10</td>
<td>F- mcrAΔ (mrr-hsdRNS-mcrBC) φ80lacZAM15 ΔlacX74deoR recA1 araD139 Δ(araleu) 7697 galU galK rpsL(StrR) endA1 nupG</td>
<td>Cloning</td>
</tr>
<tr>
<td>BL21(DE3)pLysS</td>
<td>F- ompT hsdSb (rB mB') gal dcm (DE3) pLysS (Cam')</td>
<td>Recombinant protein expression</td>
</tr>
<tr>
<td>BL21-AI</td>
<td>F- ompT hsdSb (rB mB') gal dcm araB : : T7RNAP- tetA</td>
<td>Recombinant protein expression</td>
</tr>
</tbody>
</table>
2.1.3.2 Chicken DT40 cell lines

Chicken DT40 cell lines from a gift or generated in the lab are listed in Table 2.6.

Table 2.6 DT40 cell lines used in this study

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT40 wt</td>
<td>Chicken B-cell lymphoma</td>
<td>Dr. Ciaran Morrison (NUI Galway)</td>
</tr>
<tr>
<td>$M_{sl2}^{-/-}$ #1</td>
<td>$M_{sl2}$ knockout DT40 cell line clone #1</td>
<td>This study</td>
</tr>
<tr>
<td>$M_{sl2}^{-/-}$ #2</td>
<td>$M_{sl2}$ knockout DT40 cell line clone #2</td>
<td>This study</td>
</tr>
<tr>
<td>Msl2-rescue</td>
<td>Stable re-expression of wild-type HA2F-Msl2 in $M_{sl2}^{-/-}$ #2</td>
<td>This study</td>
</tr>
<tr>
<td>Msl2-C44R</td>
<td>Stable re-expression of HA2F-Msl2 (C44R) in $M_{sl2}^{-/-}$ #2</td>
<td>This study</td>
</tr>
<tr>
<td>Msl2-ΔR</td>
<td>Stable re-expression of HA2F-Msl2 (ΔRING) in $M_{sl2}^{-/-}$ #2</td>
<td>This study</td>
</tr>
</tbody>
</table>

2.1.3.3 Mammalian cell lines

The mammalian cell lines used in this study are listed in Table 2.7.

Table 2.7 Mammalian cell lines used in this study

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>U2OS</td>
<td>Human osteosarcoma cell line</td>
<td>ATCC</td>
</tr>
<tr>
<td>HeLa CCL2</td>
<td>Human cervical adenocarcinoma</td>
<td>ATCC</td>
</tr>
</tbody>
</table>

2.1.4 Tissue culture consumables and reagents

All tissue culture plasticware were purchased from Sarstedt, Corning, Cruinn, Fisher or Sigma. Roswell Park Memorial Institute (RPMI) 1640 medium was purchased from Bio-sciences, fetal bovine serum (FBS) for DT40 cell culture from Lonza, Dulbecco’s Modified Eagle’s Medium (DMEM) and fetal bovine serum (FBS) for human cell culture from Sigma. Dimethyl Sulfoxide (DMSO), chicken serum and penicillin and streptomycin antibiotics were bought from
Sigma as well. Drugs from Sigma or Invitrogen used to screen stable cell lines are listed in Table 2.8.

**Table 2.8 Drugs used for stable cell line selection**

<table>
<thead>
<tr>
<th>Name</th>
<th>Optimal concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blasticidin (Sigma)</td>
<td>25 μg/ml</td>
</tr>
<tr>
<td>Geneticin (Invitrogen)</td>
<td>2.5 mg/ml</td>
</tr>
<tr>
<td>Puromycin (Sigma)</td>
<td>0.5 μg/ml</td>
</tr>
</tbody>
</table>

### 2.1.5 Cellular biology reagents

#### 2.1.5.1 siRNAs

Sequences of siRNAs designed for knockdown of specific genes of interest in human cells in this thesis are listed in Table 2.9. They were synthesised in Dharmacon or Sigma.

**Table 2.9 List of siRNA sequences**

<table>
<thead>
<tr>
<th>siRNAs</th>
<th>Sense target sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>siGENOMERISC-Free siRNA (Dharmacon)</td>
<td>sequence not available</td>
</tr>
<tr>
<td>hMOF #1</td>
<td>GUGAUCCAGUCUCGAGUGA[dT][dT]</td>
</tr>
<tr>
<td>hMOF #3</td>
<td>CAAGAUCACUCGCAACCAAA[dT][dT]</td>
</tr>
<tr>
<td>hMOF #4</td>
<td>GCAAAGACCAUAGAUUUAU</td>
</tr>
<tr>
<td>hMOF #5</td>
<td>GGAAGAGAUCUACCGCAA[dT][dT]</td>
</tr>
<tr>
<td>hMSL1 #1</td>
<td>GAUUGCCGUACCUUCCA[dT][dT]</td>
</tr>
<tr>
<td>hMSL1 #2</td>
<td>GUACCUUUCCACACAGA[dT][dT]</td>
</tr>
<tr>
<td>hMSL1 #3</td>
<td>AGAAAGCUUCUCAAGUGGAUAGAG</td>
</tr>
<tr>
<td>hMSL1 #4</td>
<td>GCAAAAACUGGAGCUUGGAUGAGAAG</td>
</tr>
<tr>
<td>hMSL2 ON-TARGET plus SMARTpool</td>
<td>GUGUAUGGCAGCGAAACA</td>
</tr>
<tr>
<td></td>
<td>AACAUCUAUGCCGGAAGA</td>
</tr>
<tr>
<td></td>
<td>CACCAUGCCUCCCAGAAUU</td>
</tr>
<tr>
<td></td>
<td>GUGUCAAAUUGGAGGGUAA</td>
</tr>
<tr>
<td>hMSL3 #1</td>
<td>CGGUUAGUAGAAACUUCUCAU[dT][dT]</td>
</tr>
<tr>
<td>hMSL3 #2</td>
<td>GAUCAUGUGCUUCGUGUA[dT][dT]</td>
</tr>
<tr>
<td>hMSL3 #3</td>
<td>GCAUUUGCUCAUCAUGCA[dT][dT]</td>
</tr>
<tr>
<td>hMSL3 #4</td>
<td>AAAGGUGACUUCGCUCUA[dT][dT]</td>
</tr>
</tbody>
</table>
2.1.5.2 Antibodies

Antibodies used throughout this study were mainly employed for western blot analysis and immunofluorescence microscopy. The detailed antibody information including host species, working conditions are shown in Table 2.10 and 2.11.

Table 2.10 Primary antibodies used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Host species</th>
<th>Dilution</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>alpha-tubulin (T 6074)</td>
<td>Monoclonal Mouse</td>
<td>1:10,000</td>
<td>Sigma</td>
</tr>
<tr>
<td>beta-Actin (ab8227)</td>
<td>Polyclonal Rabbit</td>
<td>1:10,000</td>
<td>Abcam</td>
</tr>
<tr>
<td>FLAG (F 1804)</td>
<td>Monoclonal Mouse</td>
<td>1:1000</td>
<td>Sigma</td>
</tr>
<tr>
<td>gamma-H2AX (05-636)</td>
<td>Monoclonal Mouse</td>
<td>1:1000</td>
<td>Millipore</td>
</tr>
<tr>
<td>GST (G 7781)</td>
<td>Monoclonal Rabbit</td>
<td>1:5000</td>
<td>Sigma</td>
</tr>
<tr>
<td>H3 (ab1791)</td>
<td>Polyclonal Rabbit</td>
<td>1:10,000</td>
<td>Abcam</td>
</tr>
<tr>
<td>H3K9me2 (#4658P)</td>
<td>Monoclonal Rabbit</td>
<td>1:1000</td>
<td>Cell Signaling</td>
</tr>
<tr>
<td>H3K9me3 (#9754S)</td>
<td>Polyclonal Rabbit</td>
<td>1:1000</td>
<td>Cell Signaling</td>
</tr>
<tr>
<td>H3K9me2 (ab3594)</td>
<td>Polyclonal Rabbit</td>
<td>1:1000</td>
<td>Abcam</td>
</tr>
<tr>
<td>H3Ser10ph (06-570)</td>
<td>Polyclonal Rabbit</td>
<td>1:1000</td>
<td>Millipore</td>
</tr>
<tr>
<td>H4K8ac (07-328)</td>
<td>Polyclonal Rabbit</td>
<td>1:1000</td>
<td>Millipore</td>
</tr>
<tr>
<td>H4K16ac (07-329)</td>
<td>Polyclonal Rabbit</td>
<td>1:1000</td>
<td>Millipore</td>
</tr>
<tr>
<td>H4K20me2 (07-367)</td>
<td>Polyclonal Rabbit</td>
<td>1:1000</td>
<td>Millipore</td>
</tr>
<tr>
<td>H4K20me3 (07-463)</td>
<td>anti serum Rabbit</td>
<td>1:1000</td>
<td>Millipore</td>
</tr>
<tr>
<td>hMOF (4G4)</td>
<td>anti serum Mouse</td>
<td>1:250</td>
<td>Dundee</td>
</tr>
<tr>
<td>hMOF (7D1)</td>
<td>anti serum Mouse</td>
<td>1:250</td>
<td>Dundee</td>
</tr>
<tr>
<td>hMOF (7G11)</td>
<td>anti serum Mouse</td>
<td>1:250</td>
<td>Dundee</td>
</tr>
<tr>
<td>hMOF (8B12)</td>
<td>anti serum Mouse</td>
<td>1:250</td>
<td>Dundee</td>
</tr>
<tr>
<td>hMOF (8E3)</td>
<td>anti serum Mouse</td>
<td>1:250</td>
<td>Dundee</td>
</tr>
<tr>
<td>hMSL2 (4F12)</td>
<td>anti serum Mouse</td>
<td>1:100</td>
<td>Dundee</td>
</tr>
<tr>
<td>hMSL2 (8A4)</td>
<td>anti serum Mouse</td>
<td>1:100</td>
<td>Dundee</td>
</tr>
<tr>
<td>hMSL2 (8D2)</td>
<td>anti serum Mouse</td>
<td>1:100</td>
<td>Dundee</td>
</tr>
<tr>
<td>MDC1 (A300-053A)</td>
<td>Polyclonal Rabbit</td>
<td>1:1000</td>
<td>Dr. Ciaran Morrison (NUI Galway)</td>
</tr>
<tr>
<td>Scc1</td>
<td>Polyclonal Rabbit</td>
<td>1:1000</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.11 Secondary antibodies used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Host species</th>
<th>Dilution WB</th>
<th>Dilution IF</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-mouse-HRP (NA931)</td>
<td>Polyclonal Sheep</td>
<td>1:10000</td>
<td></td>
<td>Amersham</td>
</tr>
<tr>
<td>Anti-rabbit-HRP (NA934)</td>
<td>Polyclonal Donkey</td>
<td>1:10000</td>
<td></td>
<td>Amersham</td>
</tr>
<tr>
<td>Anti-rabbit-FITC (711-095-152)</td>
<td>Polyclonal Donkey</td>
<td></td>
<td>1:200</td>
<td>Jackson Immuno-Research</td>
</tr>
</tbody>
</table>

2.1.6 Computer programmes

Primer Premier 5 (http://www.premierbiosoft.com/) was used for primer design. DNA plasmid maps were created using pDRAW32 (http://www.acaclone.com/). DNA sequencing data were viewed using Chromas (http://www.technelysium.com.au/chromas.html). Blast (http://www.ncbi.nlm.nih.gov/BLAST), MEGA (http://www.megasoftware.net/) and DNAMAN (http://www.lynnnon.com/) and were employed for homology alignment. Other computer softwares and programs used in this study are specified below or in the text.

2.2 Methods

2.2.1 Nucleic acid methods

2.2.1.1 RNA preparation

Total RNA was isolated from tissue culture cells using ISOLATE RNA mini kit (Bioline) according to the manufacturer’s instructions. Briefly, 5 x 10⁶ adherent or suspension cells were harvested and disrupted in lysis buffer. The released RNA was bound to the column. After washing, the RNA was eluted in 50 μl of RNase-free water and stored at -80 °C until use.

2.2.1.2 Estimation of RNA concentration

To determine RNA concentration, 1-1.5 μl of RNA was directly used to measure the absorbance at 260 nm on a NanoDrop 2000c (Thermo Scientific) spectrophotometer, as nucleic acid is detected at 260 nm, whereas protein, salt
and solvents are detected at 230 and 280 nm. High pure RNA samples with an OD 260/280 ratio of 1.8 to 2 and an OD 260/230 of 1.8 or greater will be used for further reverse transcription reactions.

2.2.1.3 cDNA synthesis

cDNA was reverse transcribed from 0.5 μg of total RNA (from Section 2.2.1.2) using the High capacity RNA-to-cDNA kit (Applied Biosystems) following the manufacturer’s instructions. The reagent conditions and programmes of reverse transcription (RT) reaction are listed below.

Table 2.12 Reverse transcription reaction conditions

<table>
<thead>
<tr>
<th>Reagent concentrations</th>
<th>RT steps</th>
<th>RT steps</th>
<th>RT steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>2× RT Buffer</td>
<td>1×</td>
<td>Reverse transcription</td>
<td>37 °C - 60 min</td>
</tr>
<tr>
<td>20× Enzyme Mix</td>
<td>1×</td>
<td>Enzyme inactivation</td>
<td>95 °C - 5 min</td>
</tr>
<tr>
<td>Total RNA</td>
<td>0.5 µg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuclease-free H2O</td>
<td>up to 10 µl</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.2.1.4 Polymerase chain reaction

Polymerase chain reaction (PCR) was performed on a TGradient or T3 Thermocycler PCR machine (Biometra) using KOD Hot Start (Novagen) or Taq (Sigma) polymerase. The PCR programs were set up according to the primers’ melting temperature and the predicted size of a PCR product. All the sequences of primers used are listed in Section 2.1.5.1. Table 2.13 and 2.14 demonstrate the PCR conditions and programmes used.

Table 2.13 PCR reaction conditions for KOD polymerase

<table>
<thead>
<tr>
<th>Reagent concentrations</th>
<th>PCR steps</th>
<th>PCR steps</th>
<th>PCR steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>10× KOD Hot Start buffer</td>
<td>1×</td>
<td>‘Hot start’</td>
<td>95 °C - 2 min</td>
</tr>
<tr>
<td>dNTPs</td>
<td>200 μM</td>
<td>Denaturation</td>
<td>95 °C - 20 sec</td>
</tr>
<tr>
<td>MgSO₄ (25 mM)</td>
<td>1 mM</td>
<td>Annealing</td>
<td>55–60 °C - 10 sec</td>
</tr>
<tr>
<td>Forward Primer (10 mM)</td>
<td>0.5 μM</td>
<td>Elongation</td>
<td>70 °C - X sec</td>
</tr>
<tr>
<td>Reverse Primer (10 mM)</td>
<td>0.5 μM</td>
<td>Final elongation</td>
<td>70 °C - 5 min</td>
</tr>
<tr>
<td>KOD Polymerase</td>
<td>1 U</td>
<td>No. of cycles</td>
<td>30–35</td>
</tr>
</tbody>
</table>
2.2.1.5 Site directed mutagenesis

Site directed mutagenesis was performed to introduce an interested point mutation in the target plasmid. Briefly, complementary primers with the desired point mutation were synthesised and used to amplify a nicked circular PCR product by using KOD Hot Start polymerase according to the modified quickchange site-directed mutagenesis protocol (Stratagene). The PCR products were then treated with Dpn I endonuclease for 2 hours, which is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template for selection of mutation-containing synthesized DNA. Finally, 5 μl of the digested product was transformed into *E. coli* competent Top10 cells to screen mutated colonies.

2.2.1.6 Generation of probes labelled with Digoxigenin (DIG) by PCR

The probes used in non-radioactive Southern hybridization were generated by PCR reaction and labelled with digoxigenin using PCR DIG Probe Synthesis Kit (Roche) according to the manufacture’s instruction. A typical example of probe labelling program is shown in Table 2.15.

Table 2.15 PCR conditions for generation of probes labelled with DIG

<table>
<thead>
<tr>
<th>Reagent concentrations</th>
<th>PCR steps</th>
<th>PCR steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x Buffer</td>
<td>1x</td>
<td>‘Hot start’</td>
</tr>
<tr>
<td>Template</td>
<td>100 pg</td>
<td>Denaturation</td>
</tr>
<tr>
<td>Primers</td>
<td>0.3 μM</td>
<td>Annealing</td>
</tr>
<tr>
<td>dNTPs</td>
<td>200 μM</td>
<td>Elongation</td>
</tr>
<tr>
<td>DIG-dUTP</td>
<td>35 or 70 μM</td>
<td>Final elongation</td>
</tr>
<tr>
<td>Enzyme</td>
<td>2.5 U</td>
<td>No. of cycles</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reagent concentrations</th>
<th>PCR steps</th>
<th>PCR steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x Buffer</td>
<td>1x</td>
<td>‘Hot start’</td>
</tr>
<tr>
<td>dNTPs</td>
<td>200 μM</td>
<td>Denaturation</td>
</tr>
<tr>
<td>MgCl2 (25 mM)</td>
<td>1.5 mM</td>
<td>Annealing</td>
</tr>
<tr>
<td>Forward Primer (10 mM)</td>
<td>0.5 μM</td>
<td>Elongation</td>
</tr>
<tr>
<td>Reverse Primer (10 mM)</td>
<td>0.5 μM</td>
<td>Final elongation</td>
</tr>
<tr>
<td>Taq Polymerase</td>
<td>1 U</td>
<td>No. of cycles</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reagent concentrations</th>
<th>PCR steps</th>
<th>PCR steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x Buffer</td>
<td>1x</td>
<td>‘Hot start’</td>
</tr>
<tr>
<td>dNTPs</td>
<td>200 μM</td>
<td>Denaturation</td>
</tr>
<tr>
<td>MgCl2 (25 mM)</td>
<td>1.5 mM</td>
<td>Annealing</td>
</tr>
<tr>
<td>Forward Primer (10 mM)</td>
<td>0.5 μM</td>
<td>Elongation</td>
</tr>
<tr>
<td>Reverse Primer (10 mM)</td>
<td>0.5 μM</td>
<td>Final elongation</td>
</tr>
<tr>
<td>Taq Polymerase</td>
<td>1 U</td>
<td>No. of cycles</td>
</tr>
</tbody>
</table>
2.2.1.7 Quantitative real-time PCR

Quantitative real-time PCR (Q-PCR) is frequently used to determine the amounts of specific transcript in RNA samples. In this study, SYBR Green Q-PCR method was chosen. SYBR Green, as a fluorescent dye, can integrate into a synthesised double-stranded DNA to monitor the accumulation of amplification products.

The amplifications were performed using the Applied Biosystem 7500 Real Time PCR System (Applied Biosystems). Briefly, cDNA reverse transcribed from RT reaction (see Section 2.2.1.3) was diluted 1:50 and amplified using primers against Msl2, Mof, Dot1l, Suv420, Gapdh, β-Actin, HoxA9 and Meis1 with Fast SYBR® Green Master Mix kit (Applied Biosystems). The reactions were run in Optical 96-well Reaction Plates with barcode (Applied Biosystem) using a tested thermal cycling program as shown in Table 2.16. At the end of the PCR reactions, samples were subjected to a melting curve analysis confirming the specificity of primers. All the reactions were performed in duplicates. The fluorescence data were subsequently analysed by using the Applied Biosystems detection system 7500 Software version 2.0.6 to determine the relative abundance of the interested genes. The relative expression levels of genes were obtained by normalising their expression to a housekeeping gene. In DT40 cells, housekeeping gene beta-Actin or Gapdh was chosen. For human cells, PPIA or GAPDH was employed. And the primer sequences are listed in Table 2.9.

<table>
<thead>
<tr>
<th>Table 2.16 Quantitative real-time PCR conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reagent concentrations</strong></td>
</tr>
<tr>
<td>2× Mastermix</td>
</tr>
<tr>
<td>Template</td>
</tr>
<tr>
<td>Forward Primer</td>
</tr>
<tr>
<td>Reverse Primer</td>
</tr>
<tr>
<td>H₂O</td>
</tr>
</tbody>
</table>

2.2.1.8 Restriction digestion of DNA

All restriction enzymes were purchased from NEB, so restriction digests of
plasmid or genomic DNA were done at optimal conditions in accordance to the NEB instructions. The digested DNA was viewed on a 1% agarose gel.

2.2.1.9 Alkaline phosphatase treatment of DNA
To prevent self-ligation after restriction digestion, the digested vector DNA was treated with Shrimp Alkaline Phosphatase (SAP) (Promega) to remove the 5’ phosphate group. 2 μg DNA samples were incubated with 1 U SAP in 1× SAP buffer at 37 °C for 2 h. For inactivation of SAP activity, samples were further incubated at 65 °C for 15 min.

2.2.1.10 Agarose gel electrophoresis
Agarose gel electrophoresis was performed according to the method described (Sambrook et al., 1989). For general use, 1% agarose gel with 0.01% SYBR Safe dye (Invitrogen) was prepared. For southern blot analysis, 0.7% agarose gel containing 0.5 μg/ml ethidium bromide was used instead. Electrophoresis was typically performed in 1x TAE buffer at 80 V using the Bio-Rad PowerPac Basic for 50 min or at 100 V for 3-4 h with the Electrophoresis Power Supply from GE Healthcare. The gels were visualised using a UV light transilluminator ChemiImager 5500 (Alpha Innotech).

2.2.1.11 DNA extraction
DNA bands were excised from the agarose gel with a scalpel blade and placed in a 1.5 ml tube. DNA was then extracted from the gel and dissolved in 30 μl sterile MilliQ water using the QIAquick Gel Extraction kit (Qiagen) according to the manufacturer’s instructions. Otherwise, DNA was directly purified by using High Pure PCR Cleanup Micro kit (Roche) to remove buffer salts, enzymes and primers.

2.2.1.12 DNA ligation
Ligation was completed in a 10 μl volume with T4 DNA ligase and buffer at
room temperature for 1 h for sticky-end ligation, or overnight for blunt-end ligation. The molecular ratios of purified vector and insert fragments varied from 1:3 to 1:10 for maximum ligation efficiency.

2.2.1.13 Preparation of competent E. coli by chemical method
The competent E. coli cells were generated by chemical method. Cells were grown in 5 ml of LB broth without any antibiotics overnight and then diluted 1:50 in 100 ml of LB broth at 37 °C with shaking to an OD600 of 0.5-0.6. The cells were incubated on ice for 5 min and spun down at 6000 g for 10 min at 4 °C. The pellet was resuspended in 50 ml ice-cold 0.1 M CaCl₂, incubated on ice for 30 minutes and pelleted again. Then, they were gently resuspended in 10 ml of ice-cold 0.1 M CaCl₂ supplemented with 15% glycerol. 50-200 μl aliquots were snap frozen in liquid nitrogen and stored at -80 °C.

2.2.1.14 E. coli transformation
The competent cells (see Section 2.2.1.13) were transformed by heat-shock at 42 °C for 45-90 seconds after incubation with DNA on ice for 15-30 min. The cells then were immediately chilled on ice for another 2 min and recovered in 1 ml of LB broth without antibiotics at 37 °C for 20-40 min with shaking. Then, cells were spread onto LB agar plates supplemented with appropriate antibiotics growing at 37 °C overnight. For selection of colonies transformed with pGEM-T Easy vector-based plasmids, LB agar plates containing ampicillin and X-gal were used for a blue/white screening (Vieira and Messing, 1982). Single white colonies on the plates were picked into LB broth with appropriate antibiotics growing overnight for plasmid extraction.

2.2.1.15 Plasmid DNA preparation
Plasmid DNA was extracted using the NucleoBond Xtra Midi Plus kit (Macherey-Nagel), Qiagen Midi Prep kit (Qiagen) or Genopure Plasmid Maxi kit (Roche) following the manufacturers’ instructions. Briefly, 3 ml, 100 ml or
200 ml of bacterial cultures was grown overnight at 37°C with shaking in the presence of selective antibiotics (50 µg/ml Ampicillin or 30 µg/ml Kanamycin (Sigma)). The next day, they were harvested and used for mini, midi or maxi plasmid isolation.

2.2.1.16 Estimation of DNA concentration
The concentration of DNA was performed by measuring the absorbance at 260 nm in cuvettes using a BioPhotometer (Eppendorf). The samples with an OD 260/280 ratio between 1.8 and 2.0 were considered as good enough for future use.

2.2.1.17 DNA sequencing
DNA samples were sent to either Cogenics (Takeley, UK) or Source BioScience (Dublin, Ireland) for commercial sequencing. The retrieved sequencing data were viewed using the Chromas program to evaluate the quality by peak profiles. Thereafter, the data with good quality were aligned with theoretical sequences provided in the database to verify the fidelity of DNA sequences of interest.

2.2.1.18 Extraction of genomic DNA from chicken DT40 cells
For southern blot analysis, ~1.5×10^6 chicken DT40 cells were disrupted in 0.5 ml of tail buffer supplemented with 0.2 mg/ml proteinase K at 37 °C overnight or at 55 °C for 3 h. The suspension lysates were vigorously rocked on a mixer at 1400 rpm for 5 min. Then, 200 µl of 6 M NaCl (saturated) were added into the lysates, shaking for another 5 min and spun down at 13000 rpm for 10 min. The supernatant was transferred to a new eppendorf tube, mixed well with 700 µl isopropanol and centrifuged at 13000 rpm for 10 min to precipitate the DNA. DNA pellet was washed once with 70% ethanol and resuspended in 100 µl milliQ H₂O.
2.2.1.19 Southern blot

10 µg genomic DNA (See Section 2.2.1.18) was used for digestion in a 50 µl reaction system at 37 °C overnight with 1 mg/ml RNase A. The digested samples were separated on a 0.7% agarose gel at 100 V for 3 h against a 1 kb DNA ladder. The gel was then washed with 0.25 M HCl and the denaturation buffer for 20 min each for depurination and denaturation of DNA to facilitate the hybridization of probe. Next, the neutralization buffer was used to soak the gel for another 20 min to equilibrate the pH of the gel. Following these, DNA fragments were transferred to a positively charged Hybond N nitrocellulose membrane (GE Healthcare) overnight at room temperature and fixed by UV irradiation (3000 J/cm²) using a UV Cross-linker (Hoefer UVC500, GE Healthcare).

For hybridisation, according to the manufacturer’s instructions (Roche), the DIG-labelled probes (see Section 2.2.1.6) needed to be boiled for 5 min to denature DNA before use. Then, the membrane pre-treated with Pre-Hyb buffer was incubated with DIG Probe diluted in Pre-Hyb buffer overnight at appropriate hybridization temperature in a hybridisation incubator (Techne). The next day, the membrane was washed with low stringency buffer at 25 °C for 2× 5 min and with high stringency buffer at 65 °C for 2× 15 min. The membrane was then blocked with non-radioactive Southern blot blocking solution for 30 min to decrease background, followed by incubation with anti-DIG-AP conjugate for another 30 min. Non-specific binding of antibody was washed out with maleic acid washing buffer. The membrane was soaked into detection buffer for 3 min to equilibrate the membrane. The membrane was put into a hybridization bag and 1:100 diluted CSPD substrate was dropped onto the membrane. Excess CSPD was squeezed out. The alkaline phosphatase (AP) cleaves CSPD to produce light that can be detected by X-ray film. The membrane was sealed in the plastic bag and incubated at 37 °C for 10 min to enhance the signal, and then exposed to an X-ray film for 5-18 h for development.
2.2.2 Protein methods

2.2.2.1 Protein sample preparation

2.2.2.1.1 Whole cell extraction

Cells were harvested at 1200 rpm for 5 min, washed once in chilled 1× PBS and then lysed in modified RIPA buffer containing protease inhibitors and phosphatase inhibitors if required on ice for 1 h. During this period, the samples were vortexed every 15 min to get better lysis effect. Samples were then centrifuged at maximum speed for 15 min at 4 °C. Each supernatant was collected and the protein sample concentration was determined by Bradford analysis (see Section 2.2.2.2).

2.2.2.1.2 Isolation of nuclear and cytoplasmic proteins

Confluent cells were harvested by centrifuging at 1200 rpm g for 5 min and washed twice with ice-cold PBS. The cells were then resuspended gently in appropriate volume of isotonic lysis buffer (50 μl/10^7 cells) and incubated on ice for 10-15 min to allow the cells to swell. This can be checked by microscopy. To the swollen cells, 10% IGEPAL CA-630 was added to a final concentration of 0.3% cells, mixed gently. Next, the cells were centrifuged at 5500 rpm for 30 sec. The supernatant (cytoplasmic protein fraction) was transferred to a chilled tube and spun down at 13000 rpm for 15 min to remove any debris. The pellet (containing the nuclei) was washed with same volume of isotonic lysis buffer again, resuspended in the appropriate volume of extraction buffer (20 μl/10^7 cells and vigorously agitated on a mixer for 15 min at 700 rpm followed by 15 min at 1400 rpm. After centrifuging at 9400 rpm for 15 min, the supernatant (nuclear protein fraction) was transferred to a chilled tube. The protein concentration was determined by Bradford analysis (see Section 2.2.2.2).

2.2.2.2 Bradford protein assay

In order to make equal loading of all protein samples for SDS-PAGE gel electrophoresis and western blot analysis, protein concentrations were
determined using the Bradford Assay. As described in the manufacturer’s instructions, 5 μl of each protein sample was mixed with 750 μl Bradford reagent (Sigma) and sit for 5 min. Then, absorbance was measured at 595 nm using a BioPhotometer (Eppendorf). Protein standards used to generate a calibration curve were obtained by stepwise dilutions of a 2 mg/ml stock solution of bovine serum albumin (BSA) (Bio-rad). Linear regression was used to determine each sample’s protein concentration.

2.2.2.3 SDS-polyacrylamide gel electrophoresis

Equal amount of protein samples were mixed with 4× Laemmli buffer containing 1.25% β-mercaptoethanol and boiled at 95 °C for 5 min to completely denature proteins. Samples were then loaded onto a discontinuous polyacrylamide gel and separated using the Mini-PROTEAN tetra cell (Bio-rad). To ensure optimal separation of proteins, 6-15% separating gel was adapted to their molecular weights. The recipe for stacking gel and different percentage of separating gels was shown in Table 2.17. Generally, samples were run at 100 V through the stacking gel and 120 V through the separating gel in 1 x running buffer. The molecular weight of proteins was determined by running a prestained protein ladder (PageRulerTM, Fermentas) in parallel.

Table 2.17 Recipe of SDS-polyacrylamide gel

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Separating gel</th>
<th>Stacking gel</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O (ml)</td>
<td>5.4 4.7 4.1 3.3 2.3</td>
<td>3.05</td>
</tr>
<tr>
<td>30% Bis-Acrylamide (ml)</td>
<td>2 2.7 3.3 4.1 5.1</td>
<td>0.65</td>
</tr>
<tr>
<td>1.5 M Tris-HCl pH 8.8 (ml)</td>
<td>2.5</td>
<td>-</td>
</tr>
<tr>
<td>0.5 M Tris-HCl pH6.8 (ml)</td>
<td>-</td>
<td>1.25</td>
</tr>
<tr>
<td>10% SDS (μl)</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>10% APS (μl)</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>TEMED (μl)</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

49
2.2.2.4 Coomassie staining
Proteins were resolved using an acrylamide gel (see Section 2.3.4). Gels were stained with Coomassie Blue R250 for 30 min at 60 °C with gentle agitation, then destained overnight at room temperature with several rounds of destain solution.

2.2.2.5 Transfer of proteins to membrane
Proteins separated by SDS-PAGE were transferred to nitrocellulose membranes (Watman) or PVDF membranes (GE Healthcare) using semi-dry or wet transfer method in 1× transfer buffer system. Depend on the target protein size, semi-dry transfer was carried out at 25 V for 45-90 min at room temperature using the Trans-Blot SD Semi-dry Transfer Cell (Bio-rad), the wet transfer was performed at 100 V for 60-90 min at 4 °C using the Mini Trans-Blot Cell (Bio-rad). Following the transfer, the membrane was stained with Ponceau S for several minutes to ensure the quality of protein transfer after one wash with water.

2.2.2.6 Immunoblotting
Post-transferred membranes (see Section 2.2.2.5) were blocked with 2-5% milk or BSA in 1× PBST at room temperature for 30 min with gentle agitation on a shaker to minimise non specific binding. Subsequently, membranes were incubated with properly diluted primary antibodies (Table 2.10) at room temperature for 2 h or overnight at 4 °C with gentle shaking. The excess antibodies were washed out with 3 fresh changes of 1× PBST at room temperature. The membrane was then incubated with the secondary antibody at room temperature for 1 h with gentle agitation. After another 3× 5 min washes, the proteins of interest were detected with ECL plus Western Blotting detection reagents (GE Healthcare) according to the manufacturer’s instructions. The enhanced chemiluminescence signal was exposed to an X-ray film (Konica Minolta) and developed using a CP 1000 developing machine (AGFA) or digitally captured using a LAS-3000 Imager (Fujifilm). The intensity of Western
blot bands was quantified using the ImageJ software Version 1.43q.

**2.2.3 Cell biology techniques**

**2.2.3.1 Maintenance of cells**

Chicken DT40 cells were cultured as described in RPMI 1640 (Bio-sciences), supplemented with 10% fetal calf serum (Lonza), 1% chicken serum (Sigma) and 1% penicillin/streptomycin (Sigma) at 39.5 °C with 5% CO\(_2\) in humidified conditions (Takata et al., 1998). Cells grow in suspension. When confluent, the cell density is up to 1×10\(^6\) cells/ml.

Human adherent cells were cultured in DMEM supplemented with 10% Fetal Bovine Serum (Sigma) at 37 °C with 5% CO\(_2\) under 95% humidity. Cells were digested with trypsin and split every 2-3 days when they were up to 90% confluent.

**2.2.3.2 Freezing and thawing cells**

For DT40 cells, 5×10\(^6\) cells were harvested by centrifugation at 1200 rpm for 5 min. The pellet was resuspended in 400 µl freezing medium (FBS, 10% DMSO) and transferred to a cryovial. For human cells, 3×10\(^6\) cells were pelleted at 1200 rpm for 5 minutes, resuspended in 1 ml freezing medium (DMEM, 10% DMSO, 20% FBS), and transferred to a cryovial. The cryovials were then put into a Cryo 1 °C freezing container (Nalgene) and sit at -80 °C overnight. Then, cells could be kept at -80 °C for up to a year and transferred to liquid nitrogen for longer-term storage.

To wake up cells, frozen cells were quickly thawed at 37 °C, gently transferred to Falcon tubes containing 5 ml growth medium and pelleted as described above. The cell pellets were resuspended with prewarmed medium and plated out.

**2.2.3.3 Cell proliferation analysis**

Cells were counted using a haemocytometer as described in the manufacturer’s protocol. For DT40 cells, each cell line was diluted into 24-well plates in
triplicate with equal cell densities ($5 \times 10^4$ cells/ml). The cells were cultured and counted every 24 hours up to 96 hours. When cells were growing up to 50-70% confluent, they were diluted back to the initial cell density for culture. The dilution factor was taken to calculate the cell number and plot growth curves.

2.2.3.4 Cell transfection

2.2.3.4.1 Stable transfection of DT40 cells

To generate stable DT40 cell lines, electroporation method was employed here. For each transfection, $1 \times 10^7$ cells were harvested and washed once with $1 \times$ PBS. The cell pellet was then resuspended in 500 µl $1 \times$ PBS and mixed with 15-20 µg of linearised plasmid DNA in a 0.4 cm gap electroporation cuvette on ice for 10 min. Electroporation was performed at either 550 V/25 µF or 300 V/600 µF using a Gene Pulser electroporation apparatus (Bio-rad). Thereafter, the cuvette was put back on ice for another 10 min, and the electroporated cells was transferred into a 10 cm dish containing 20 ml of prewarmed medium for overnight culturing. The next day, the cells were added one volume of prewarmed medium containing appropriate amount of selection drug and seeded into 96-well plates with 100 µl/well. The plates were kept in the incubator for 7-10 days to wait visible single colonies coming up. Each single colony was further picked up into 24-well plates until growing confluent. The positive colonies were identified by southern blot or western blot analysis.

2.2.3.4.2 Transient transfection of human cells

Human cells were transiently transfected using either Lipofectamine 2000 or oligofectamine (Invitrogen). Plasmids or siRNAs used for transfection are presented in Table 2.4 and 2.9. Small interfering RNAs (siRNAs) were used to knock down the expression of interested genes. While plasmids were transfected to over-express the interested proteins in cells. The day before transfection, appropriate numbers of cells were plated into proper cell culture dishes or plates following the manufacturer’s guidelines. For
siRNA transfection, oligofectamine transfection reagent was chosen and cells need to grow 30-40% confluency. To transfect plasmids, lipofectamine 2000 was used instead and cells up to 80-90% confluency. siRNA, plasmids or transfection reagents was separately diluted in the same volume of reduced serum medium Opti-MEM (Invitrogen) and incubated at room temperature for 5 min. Then diluted siRNA and oligofectamine were mixed together for another 20 min. The same was done to the plasmid and Lipofectamine 2000 mixture. Lipofectamine 2000 transfection complex was directly dropped onto cells for 4 h incubation at 37 ºC before being replaced with 2 ml of DMEM medium. For siRNA transfection, cells were washed with Opti-MEM once before the transfection mixture was added. After incubation for 4-6 h, 3× DMEM medium (DMEM, 3× FBS, 3× L-glutamine) was added. The cells were further growing until 48-72 h for analysis.

2.2.3.5 Clonogenic survival
A methylcellulose colony formation assay was performed as described previously (Takata et al., 1998). Briefly, sensitivity to γ-radiation was measured by exposing cells plated in medium containing 1.5% (w/v) methylcellulose to a 137Cs source at 23.5 Gy/min (Mainance Engineering). To measure the sensitivity of cells to olaparib (PARP inhibitor) or ICRF-193 (Toposiomerase II inhibitor), cells were continuously exposed to various concentrations of Olaparib mixed in methylcellulose medium. The survival of cells to MMS were measured by culturing them in 1 ml of complete medium with increasing doses of MMS at 39.5°C for 2 h followed by serial dilution into methylcellulose medium. Colonies formed by surviving cells were counted 7-10 days later. Survival is expressed as a percentage, using untreated cells as the 100% value.

2.2.3.6 Flow cytometry
For analysis of cell cycle profiles, 5× 10⁶ DT40 cells were centrifuged at 1200 rpm for 5 min and resuspended in 1 ml of 70% ethanol in 1× PBS for fixation at
-20 °C. For propidium iodide (PI) profiles, cells were washed in 1× PBS and incubated in 1× PBS containing 40 μg/ml PI and 200 μg/ml RNaseA overnight at 4 °C in the dark. For phospho-Ser10 Histone H3 (H3S10ph) FACS, fixed cells were washed in 1× PBS and resuspended in 1 ml of 1× PBS containing 0.25% Triton X-100. After 15 min incubation on ice, cells were pelleted and then incubated with 1% anti-H3S10ph antibody in 1× PBS with 1% BSA for 2 h with shaking at room temperature followed by washing and 30 min incubation with FITC anti-rabbit antibody at 1:30 in 1× PBS with 1% BSA at room temperature in the dark. Finally, cells were resuspended in PI/RNaseA solution as above. Flow cytometry was performed on a FACS Canto (Becton Dickinson) and analysed using BD FACS Diva Software version 6.1.2.
Chapter 3 Generation of DT40 Msl2<sup>−/−</sup> and Msl2<sup>−/−</sup>-rescue cell lines

3.1 Introduction

The MSL2 gene was first identified in the fruitfly Drosophila melanogaster. The encoded MSL2 protein is a critical factor in the process of dosage compensation (Kelley et al., 1995; Villa et al., 2012). It is evolutionarily conserved from fly to human (Li et al., 2009; Smith et al., 2005).

Kuroda and colleagues cloned MSL2 in 1998, and first characterized the MSL2 protein (Copps et al., 1998). It contains three structured domains; an N-terminal RING finger domain, a C-terminal cysteine-rich (CXC) domain and a basic, proline-rich patch (Pro/Bas patch) (Copps et al., 1998; Fauth et al., 2010). MSL2’s RING finger region was first suggested to interact with MSL1 (Copps et al., 1998). Then, MSL1/2 proteins were thought to form a heterotetramer via their putative coiled-coil regions (Li et al., 2005; Scott et al., 2000; Wu et al., 2011). Recently, Hallacli et al proposed that two molecules of MSL2 can only interact with the preformed MSL1 homodimer (Hallacli et al., 2012). The MSL1 homodimerization is mediated by its coiled-coil region. MSL2 interacts with MSL1 via its coiled coil region. Both human and Drosophila CXC domains of MSL2 are important for DNA binding, although most amino acids between them are not conserved (Fauth et al., 2010). For the proline-rich patch motif, there are currently no biological findings linked to it.

Of the three domains, the RING finger is most studied and often possesses E3 ubiquitin ligase activity. In 2009, hMSL2 was first reported as a member of RING finger E3 ubiquitin ligase family (Kruse and Gu, 2009).

However, relatively little is known about the human protein. So to better explore the function of hMSL2, we first took a loss of function approach using siRNAs in cultured cells. Unfortunately, the knockdown efficiency even with siRNA smart pool was not satisfactory; three rounds of siRNA treatment were needed to deplete hMSL2. This regime takes a prolonged time, stresses the cells, and
increases the chances of non-specific effects of the treatment. Also, the depletion does not always achieve acceptable reduced levels of protein. Furthermore, we now suspect that the protein is expressed at low basal levels, so that incomplete knockdown may not manifest a true phenotype. We therefore decided to carry out a complete knockout of this gene. DT40 cells are widely used for efficient gene disruptions due to its high homologous recombination activity. So to completely ablate Msl2 protein, we attempted to disrupt the Msl2 gene locus in chicken DT40 cells.

3.2 Sequence analysis of Msl2 protein

The MSL2 protein has been found in a variety of organisms, including Xenopus, chicken, mouse and human, since it was characterized in Drosophila. So far, at least two isoforms of MSL2 have been found in Drosophila, Xenopus and human in the NCBI database. For chicken and mouse, to date, we could only find one isoform. We used the longer isoforms of MSL2 from these species chosen to build a phylogenetic tree. The tree showed that MSL2 proteins from these five selected species are descended from a common ancestor (Figure 3.1A). The chicken protein Msl2 is in a central position along the evolutionary history. But according to the homology values, the chicken Msl2 showed a higher similarity to the mammalian homologs than to those from non-vertebrate organisms. The homology between chicken Msl2 and human MSL2 is up to 84%, which is only 19% in comparison with Drosophila. The predicted secondary architectures showed that all MSL2 homologs consist of an N-terminal RING finger domain and a C-terminal CXC DNA binding domain (Figure 3.1B).
Chapter 3 Generation of DT40 $Msl2^{-/-}$ and $Msl2^{-/-}$-rescue cell lines

Figure 3.1 Analysis of Msl2 protein sequences from human, mouse, chicken, Xenopus and Drosophila. A) Phylogenetic tree of MSL2 protein was constructed with neighbor-joining method. The taxa joined together in the tree indicate that they descended from a common ancestor. Numerical values show the homology between them. B) Predicted secondary structure of MSL2 protein.

To further validate the use of the chicken DT40 system we aligned chicken Msl2 with human MSL2 isoforms 1 and 2. As shown in Figure 3.2, the chicken Msl2 is highly conserved with its human homolog, sharing 83% identity and 90% similarity with human isoform 1 across the full length of their sequences. The sequence identity between their N-terminal RING domains is even up to 98%. The human isoform 2 lacks most RING finger sequence, implicating that this isoform probably has no E3 ligase function. This isoform has not been studied in human cells, and its relevance is not known. All isoforms contain the conserved C-terminal CXC domain, and have 100% identity.


Figure 3.2 Multiple alignment of the sequence of Msl2/hMSL2 between chicken and human species. The conserved RING domain cysteine and histidine residues are highlighted in red and bold. The RING domain sequences are shaded in yellow. The C-terminal CXC domain sequences are shaded in green. Human-L refers to human MSL2 long isoform (isoform 1) sequence and Human-S refers to human MSL2 short isoform (isoform 2).

3.3 Mapping of the chicken Msl2 genomic locus

The above bioinformatic analysis revealed that the chicken Msl2 protein is so highly conserved to its human homolog, so it is possible for us to better understand the biological function of hMSL2 via studying its chicken homolog. In this study, we were going to investigate Msl2’s function with loss of function
method in DT40 cell line. To perform integrated gene targeting, the detailed Msl2 gene locus information is required, as this is essential in designing the appropriate knockout strategy and the relative targeting vectors for Msl2 disruption.

Based on information provided on the NCBI database, the chicken Msl2 gene is localised on chromosome 9 and spans about 14.5 kb across the genome. Being on chromosome 9 means there are two alleles of the gene in DT40 cells. The chicken Msl2 gene comprises only two exons, and is predicted to encode a 579 amino acid protein (Figure 3.2). While the human MSL2 gene is located in chromosome 3 (3q22.3), contains two exons and finally translated into two isoforms due to an alternate 5' end exon. The longer isoform 1 encodes 577 amino acids, while the short isoform 2 encodes 503 amino acids (Figure 3.2). The mouse Msl2 gene is also located on chromosome 9 and encodes 577 amino acids.

Although Msl2 localises to different chromosome in chicken, mouse and human, the neighboring genes surrounding Msl2 are the same. The syntenic gene map of Msl2 from chicken, human and mouse is shown in Figure 3.3. The Ppp2r3a, Pccb, Stag1 and Slc35g2/Tmem22 genes are located in similar flanking positions relative to Msl2 in all three species. Given the above, we concluded that chicken, mouse and human MSL2 are in syntenic genomic regions.

Although Msl2 localises to different chromosome in chicken, mouse and human, the neighboring genes surrounding Msl2 are the same. The syntenic gene map of Msl2 from chicken, human and mouse is shown in Figure 3.3. The Ppp2r3a, Pccb, Stag1 and Slc35g2/Tmem22 genes are located in similar flanking positions relative to Msl2 in all three species. Given the above, we concluded that chicken, mouse and human MSL2 are in syntenic genomic regions.

![Figure 3.3 Syntenic gene profile of Msl2 loci from chicken, human and mouse.](image)

The flanking genes from either side of MSL2 were compared between chicken, human and mouse species. The gene information is from NCBI database. The homologous genes shown in chicken, human and mouse are highlighted with green and those only in human and mouse with blue.
3.4 Cloning chicken Msl2

Both the gene synten and protein sequences are conserved between human and chicken. Next, we wanted to confirm the expression of chicken Msl2 in DT40 cells. The predicted mRNA sequence of chicken Msl2 is accessed in the NCBI database (Accession: XM_426675). In order to amplify chicken Msl2 cDNA, the mRNA pool was firstly extracted from chicken DT40 cells and subjected to reverse transcription. PCR was then carried out to amplify full length Msl2 cDNA using specific primers. The primers were designed according to mRNA sequences obtained from the NCBI database. The PCR products were then run on a 1% agarose gel, two bands were detected by UV exposure. As shown in Figure 3.2, the size of the longer band was between 1.5 and 2.0 Kb, which is consistent with the predicted size of 1740 bp published in the NCBI database. While the shorter one was only approximately 900 bp, we questioned if a shorter isoform of Msl2 also exists in chicken. Subsequently, the two PCR products were cloned into pGMT-Easy vector for DNA sequencing. The alignment of the sequencing data from longer PCR product and the predicted sequence provided in the NCBI database showed 100% identity, suggesting that the Msl2 cDNA was successfully cloned. But according to the sequencing result, the shorter PCR product was just some non-specific band.

![Figure 3.4 Cloning of chicken Msl2 cDNA.](image)

The Msl2 PCR product amplified using Msl2-specific primers was detected on the agarose gel by SYBR Green staining. M represents the 1 kb molecular weight marker.
3.5 Generation of DT40 Msl2<sup>−/−</sup> and Msl2<sup>−/−</sup>-rescue cell lines

3.5.1 Construction of DT40 Msl2<sup>−/−</sup> cell lines

The chicken Msl2 gene spans approximately 14.5 kb and comprises two exons. As can be seen in Figure 3.5, 92% of the protein (from amino acid 48 to 579) is encoded by exon 2. This includes the conserved N-terminal RING domain, C-terminal CXC DNA binding domain and part of Msl1 binding site. We therefore rationalized that targeting this second exon for depletion would result in a functionally null product. So the knockout strategy was designed as shown in Figure 3.5. Exon 2 was replaced with selection marker cassettes by homologous recombination. A 5’ probe external to the targeting construct was chosen for Southern blotting analysis to monitor gene-targeting.

![Figure 3.5 Strategy for creation of DT40 Msl2 knockout cell line.](image)

To test the probe’s specificity in Southern blot, the genomic DNA from DT40 wild-type cells was digested with several restriction endonucleases and hybridised with the probe. For each selected enzyme, the predicted DNA band size is shown in Figure 3.6A. Figure 3.6B is a representative southern blot result of the probe’s specificity test. For any of these restriction enzymes, only one clean band with predicted size was detected, which indicated that this probe is suitable for further screening of targeted clones.
Chapter 3 Generation of DT40 Msl2\textsuperscript{\textminus}\textendash\textminus and Msl2\textsuperscript{\textminus}\textendash\textendash -rescue cell lines

Figure 3.6 Testing of the nonradiolabelled 5’ probe by Southern blot hybridization. The genomic DNA from DT40 wild-type cells was digested with the indicated restriction endonucleases and probed with 5’ probe. A) Theoretical size of digested DNA fragment in Southern Blot analysis. B) A representative southern blot result using the relative restriction enzymes.

For targeting the first allele of Msl2, either an Msl2-Pur or Msl2-Bla disruption construct was separately electroporated into DT40 wild-type cells. The targeting efficiency of the two plasmids was different; 20% for Msl2-Pur construct and only 7.7% for Msl2-Bla cassette (Table 3.1). So to increase the number of positive clones from the second round of targeting, one heterozygous clone with blasticidin resistance was used as a parent and electroporated with Msl2-Pur construct to disrupt the remaining allele.

Table 3.1 Targeting frequencies of the targeting constructs

<table>
<thead>
<tr>
<th></th>
<th>Puromycin</th>
<th>Blasticidin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total clone number</td>
<td>85</td>
<td>52</td>
</tr>
<tr>
<td>Positive clone number</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>Recombination efficiency (%)</td>
<td>20%</td>
<td>7.7%</td>
</tr>
</tbody>
</table>

Successful targeting of both alleles was confirmed by Southern blotting. Finally, two Msl2\textsuperscript{\textendash\textendash} clones, Msl2\textsuperscript{\textendash\textendash} #1 and Msl2\textsuperscript{\textendash\textendash} #2, were screened from 72 clones (Figure. 3.7A). The targeting frequency was only 2.8% this time. The loss of Msl2 was also confirmed by reverse transcription PCR (RT-PCR) (Figure.
3.7B).

Figure 3.7 Identification of Msl2 knockout cell lines by Southern blot and RT-PCR. A) Southern blot confirmation of targeting. Genomic DNA from wild-type DT40 cells and clones after the first (+/-) or second (-/-) round of targeting were analyzed by Southern hybridization with the probe. The wild-type allele is detected at 4.2 kb, and the target allele is detected at 8.6 or 9 kb. B) RT-PCR analysis of expression levels of β-Actin and Msl2 mRNA in DT40 wild-type and Msl2^-^- cell lines.

3.5.2 Generation of Msl2^-^- rescue cell lines

To substantiate any phenotypes found in Msl2 knockout cells, we wanted to generate a rescue DT40 stable cell line. We used one of the above Msl2 knockout clones as recipient for re-expression of Msl2. In order to monitor this rescue protein, we included a Haemagglutinin-FLAG-FLAG-tag (HA2F) at the N-terminus of the rescue protein. We also wanted to create additional stable cell lines, with the same HA2F-tag as supplementary controls, and to study the Msl2 RING domain.

We constructed the pCDNA3.1-based plasmids: HA2F-Msl2 (rescue), encoding the FLAG-tagged wild-type Msl2; HA2F-Msl2 (C44R), encoding a RING domain cysteine to arginine mutation; and HA2F-Msl2 (ΔRING), encoding a truncated Msl2 protein lacking the RING domain-containing N-terminal 107 amino acids.

The plasmids were initially transfected into U2OS cells to test protein expression. After 48h of transfection, cells were harvested for whole cell lysate extraction. As shown in Figure 3.8, all over-expressed wild-type and mutant
Msl2 proteins were detected by western blot using mouse anti-FLAG antibody. Compared to the wild-type and truncation of Msl2, the point mutant was expressed at a much lower level. Extra small bands were also detected in cell lysates with expression of the full length and truncation of Msl2 by the antibody, indicating there is probably some degree of degradation.

Figure 3.8 Expression of full length or mutant Msl2 in U2OS cells. The plasmid pCDNA3.1-HA2F-Msl2 (rescue), pCDNA3.1-HA2F-Msl2 (ΔRING), pCDNA3.1-HA2F-Msl2 (C44R) or pCDNA3.1-HA2F vector (Mock) was separately transfected into U2OS cells. 48h after transfection, whole cell lysates from cells were subjected to western blot analysis to detect expression of full length and mutant Msl2 proteins using mouse anti-FLAG antibody. The apparent size of Msl2 (rescue) and Msl2 (C44R) is ~ 75 kD. It is ~ 65 kD for Msl2 (ΔR).

The plasmids were then separately electroporated into Msl2−/− #2 cells. Cells were screened using geneticin selection. For the Msl2-rescue cell line, 23 separate clones were selected for western blot analysis, and 15 positive clones were identified. For the Msl2-ΔR cell line, four positive clones were obtained from 32 total clones, while for the Msl2-C44R cell line, 17 positive clones from 23 separate clones were screened.

For further analysis of the Msl2-Rescue cell line, the nuclear extracts from six Msl2-rescue clones were prepared for western blot analysis (Figure 3.9A). Compared to the wild-type and knockout parent cells, the re-expression of Msl2 were successfully detected in all six clones by anti-FLAG antibody. The expression levels between different clones were not the same when they were normalized to those of β-Actin. Also at the same time, the change of H4K16ac
was characterized, as we have known that this histone modification was down-regulated in \( Msl2 \) knockout cell lines (See Chapter 4). Histone modification levels were normalized to histone H3 expression levels. Figure 3.9B shows the quantification of the levels of HA2F-Msl2 and H4K16ac in selected stable cell clones. Intriguingly, the expression levels of Msl2 exhibits some negative correlation with the histone modification. In the \( Msl2 \)-Rescue clone 17, Msl2 expression is the lowest compared to the other five, however its mRNA level is around 3.5 folds to that of wild type (Figure 3.9C). The affected H4K16ac was restored to almost wild-type levels. So this clone was picked up for the following experiments.

**Figure 3.9 Screening stable \( Msl2 \) rescue DT40 clones.** Six \( Msl2 \) rescue clones (C2, C3, C5, C12, C13 and C17) were screened. A) Western blot analysis of Msl2 expression at protein level using the anti-FLAG antibody. B) Quantification of (A). Levels of H4K16ac in selected stable cell clones were quantified and graphed relative to those in wild-type cells after normalization to H3 levels. Expression levels of HA2F-Msl2 are normalized to \( \beta \)-Actin, and graphed relative to clone 17. C) Q-PCR derived expression levels of \( Gapdh \) and \( Msl2 \) mRNA in DT40 wild-type, \( Msl2^{+/} \) and \( Msl2 \)-rescue cell lines. Expression is normalized to \( \beta \)-Actin and compared to wild-type levels. Error bars represent standard deviation (n=3).
For the two Msl2 mutant-expressing cell lines (ΔR and C44R), Figure 3.10 is a representative result showing the mutant HA2F-Msl2 expression at both protein and transcript level.

As can be seen in Figure 3.10A, we have four positive Msl2-ΔR cell clones; 4, 7, 11 and 23. The predicted size of the RING-deleted Msl2 (Msl2-ΔR) protein is around 65 kD. Interestingly, there were three bands shown in this region. One possible explanation is that the top band is the full length of Msl2-ΔR protein, and the small bands represented the degraded protein as some extra smaller bands also detected between 37 and 50 kD. The other possibility is that there exists some kind of posttranslational modification. The transcript level of truncated Msl2 was also evaluated. We found that the level of mRNA expressed from the integrated plasmid to be much higher than endogenous levels in wild-type cells (Figure 3.10B). As a control, we also looked at Mof mRNA expression. Compared to wild type cells, the Mof transcripts were not affected in the various clones. Considering both the transcription and protein levels of Msl2, finally, the Msl2-ΔR clone 4 was chosen for further use.

Eight positive Msl2-C44R clones were analysed by western blot (Figure 3.10C). The point mutated Msl2 protein was specifically detected in each clone with the size 75 kD. The clone 4, 6, 12 and 14 were further chosen for Q-PCR analysis to compare the Msl2 expression at mRNA level. Similar to the Msl2-ΔR clones, the expression of Mof were almost the same in these clones, while the Msl2 expression varied from clone to clone. Msl2-C44R clone 4 was chosen for the future use.
Figure 3.10 Screening truncated or point mutated Msl2 re-expressed stable DT40 clones. A) & C) Western blot analysis of truncated or point mutated Msl2 expression in DT40 Msl2-ΔR and Msl2-C44R stable clones by the anti-FLAG antibody. B) & D) Q-PCR derived expression levels of Gapdh, Msl2 and Mof mRNA in the relative mutant clones, including the wild-type, Msl2-/-, Msl2-rescue cells as controls. Expression is normalized to β-Actin and compared to wild-type levels.

3.6 Localization of the chicken Msl2 protein

MSL2 has been shown to localize to the nucleus in human and Drosophila by immunofluorescence data (Copps et al., 1998; Kruse and Gu, 2009). We were interested to see whether chicken Msl2 displays the same localization in DT40 cells. To address this, the Msl2-rescue cell line was used. We detected the
HA2F-Msl2 fusion protein using an anti-FLAG antibody (Figure 3.11A). Unfortunately, a non-specific band was also present just above the target Msl2 band. Due to the interference of this non-specific signal, it is hard to distinguish Msl2 localization in DT40 cells by indirect immunofluorescence. We therefore tried to analyze the localization of Msl2 by western blotting using nuclear and cytoplasmic fractions of Msl2-rescue cells (Figure 3.11B). HA2F-Msl2 protein was found to exist mostly in the nucleus, while the non-specific band appeared in cytoplasmic fraction. Here, chicken Scc1 was included as a nuclear marker and beta-Actin as a cytoplasmic marker.

**Figure 3.11 Nuclear localization of chicken Msl2 protein in DT40 cells.** A) Whole cell lysates from DT40 Msl2−/− #2 and Msl2-rescue cells were analysed by western blot. B) Nuclear and cytoplasmic fractions from DT40 Msl2−/− #2 and Msl2-rescue cells were analysed by western blot. Re-expressed Flag-tagged Msl2 protein was detected using anti-FLAG antibody. Scc1 and beta-Actin were used as nuclear and cytoplasmic indicators. N represents nuclear fraction and C represents cytoplasmic fraction. The predicted size of FLAG-Msl2 is ~75 kD, 110 kD for Scc1, and 43 kD for beta-Actin. Non-specific band is indicated with an open arrowhead. HA2F-Msl2 is indicated with an arrowhead.

### 3.7 Discussion

Taking a loss of function approach in human cells using siRNA was not feasible for the study of hMSL2. We decided to use the genetically tractable chicken DT40 system instead. The DT40 system is a widely acceptable model for the functional study of a gene of interest using reverse genetic analysis. Initial examination of the chicken Msl2 protein revealed a significant homology to the human and mouse homologs, with a relatively lower score to *Drosophila* Msl2
Chapter 3 Generation of DT40 Msl2<sup>−/−</sup> and Msl2<sup>−/−</sup>-rescue cell lines

homolog. The domains with obvious structure and predicted function; the RING finger domain and the CXC DNA binding domain, show almost 100% identity or similarity between chicken and human MSL2 protein (Figure 3.2). Furthermore, the chicken, human and mouse genomes show synteny around the Msl2/hMSL2 locus. Msl2 cDNA was successfully cloned in this study, which indicates that the chicken Msl2 gene is transcribed in these female DT40 cells. Using the chicken Msl2 gene locus map, we devised a strategy to disrupt the Msl2 gene by replacement of exon 2 with selective resistant cassettes. Finally, the Msl2<sup>−/−</sup> cell lines were generated and two separate clones were used to examine any functional defect with the loss of Msl2.

Three additional stable cell lines were also made using the Msl2<sup>−/−</sup> #2 cell line, including the Msl2-rescue, Msl2-ΔR and Msl2-C44R cell lines. The Msl2-rescue was made to ensure that any phenotypes observed in the knock-out cell lines are due to loss of Msl2, and should be rescued by re-expression of the full length wild-type Msl2 protein in this line. The Msl2-ΔR cell line expresses a truncated Msl2 protein lacking the N-terminal 1-107 amino acids. The same amino acid truncated human homolog has been shown to lose its E3 ubiquitination activity and the interaction with hMSL1, but still can localise to the nucleus when expressed exogenously (Hallacli et al., 2012; Kruse and Gu, 2009). So the use of this cell line would help to determine whether the phenotypes observed in the Msl2<sup>−/−</sup> cells are due to Msl2’s E3 ligase activity. The Msl2-C44R cell line was generated for the same objective. It is known that the C3HC4 motif of hMSL2 is chelated two Zn atoms to exert its E3 ligase function (Hallacli et al., 2012). This motif consists of the conserved cysteines; 44, 47, 67, and 70 and cysteine 62, 81, 84 and histidine 64 respectively. So the Msl2-C44R cell line, where cysteine at position 44 is replaced by arginine, should lack this E3 ligase activity. This cell line should also help us to determine if any phenotypes observed in the knockout are due to Msl2 E3 ligase activity.

With the rescue cells, we first tested the expression of Msl2 using the Q-PCR and western blot analysis. The mRNA amount is only ~3.5 fold more than DT40
wild-type cells. Using this cell line, we can detect the Msl2 protein expression with an anti-FLAG antibody and monitor the endogenous change in Msl2 expression. Similar to its human homolog, chicken Msl2 also localises in nucleus and probably interacts with the other Msl homologs to form a stable complex. With the nuclear extraction method, Msl2 protein can be easily separated from the non-specific band on western blot, which will facilitate the following experiments, such as immunoprecipitation.
Chapter 4 Functional analysis of $Msl2^{+/+}$ and $Msl2^{−/−}$-rescue DT40 cells

4.1 Introduction

hMSL2 is a component of the human MSL complex. It contains an N-terminal RING finger domain and C-terminal CXC domain (Marin, 2003; Smith et al., 2005). It was first reported as an E3 ligase by the Gu group in 2009, ubiquitylating p53 at two lysine sites 351 and 357, which results in translocation of p53 from the nucleus to the cytoplasm. This subcellular change in the localization of p53 was believed to mediate apoptosis (Kruse and Gu, 2009). Then in 2011, Dou’s group published data that showed hMSL2 could monoubiquitylate histone H2B. Loss of H2B-K34 monoubiquitination in hMSL2 knockdown cells caused down regulation of H3K4me3 and H3K79me2 and thus, further influenced transcription regulation (Wu et al., 2011). It was recently shown that Drosophila MSL2 can ubiquitylate not only itself, but also other members of the MSL complex, such as MSL1, MSL3 and MOF (Villa et al., 2012). But how hMSL2 is regulated, its E1 and E2 partners, and if it has other substrates is not known.

In the human MSL complex, the protein hMOF has been shown to play important roles in different cellular processes. As mentioned previously, hMOF globally mediates H4K16ac (Gupta et al., 2008; Li et al., 2010; Smith et al., 2005; Taipale et al., 2005) and is involved in transcription regulation (Kapoor-Vazirani et al., 2008; Li et al., 2009), cell cycle progression (Smith et al., 2005; Taipale et al., 2005) and the DNA damage response (Sharma et al., 2010; Taipale et al., 2005). Moreover, the hMSL1 protein, directly interacting with hMSL2, was shown to interact with DNA damage mediator 53BP1 (Gironella et al., 2009). Knockdown of mouse MSL1, or MOF induced less mediator of DNA damage checkpoint 1 (MDC1) foci at the DNA damage sites (Li et al., 2010). Given all of these observations, we questioned if hMSL2
protein, similar to its hMSL associates, functions in these biological events, especially in DNA damage response.

As mentioned previously, a loss of function approach in human cells was not feasible. Therefore, to address the above questions, we took advantage of the DT40 Msl2 knockout and rescue cells generated in the lab.

4.2 Chicken Msl2 is not essential for DT40 viability

Our first observation was that the Msl2<sup>−/−</sup> cells were viable, with cell morphology visibly similar to wild-type (data not shown), demonstrating that Msl2 is not an essential gene in DT40 cells. We then examined the proliferative ability of Msl2<sup>−/−</sup> cells. The growth assay was done by comparison of DT40 wild-type, Msl2<sup>−/−</sup> and Msl2-rescue cells. We found that cells lacking Msl2 proliferate more slowly than wild-type cells (Figure 4.1A). Wild-type cells proliferated with a doubling time of 8.05 hours whereas Msl2<sup>−/−</sup> #1 and #2 took 9.18 and 9.62 hours respectively, a delay of approximately 20%. The rescue cells displayed a recovery of this delay, having a doubling time of 7.95 hours.

Then, we further investigated the growth rate of Msl2-ΔR cells compared to DT40 wild-type, Msl2<sup>−/−</sup> #2 and Msl2-rescue cells. As shown in Figure 4.1B, similar to Msl2<sup>−/−</sup> cells, the Msl2-ΔR cells grow slower than wild-type cells. The re-expression of truncated Msl2 was unable to restore the growth defect, and it has an even more pronounced growth defect, with a doubling time of 10.6 hours. The exact cause of this growth delay in Msl2<sup>−/−</sup> cells is unclear.
Chapter 4 Functional analysis of Msl2\(^{-/}\) and Msl2\(^{-/}\)-rescue DT40 cells

Figure 4.1 Proliferation analysis of Wild-type, Msl2\(^{-/}\), Msl2-rescue and Msl2-\(\Delta\)R cell lines. 10\(^4\) cells were plated and counted at times indicted. Every 24 h, 10\(^4\) cells were replated for culture. The dilution factor was taken to calculate the cell number and plot growth curves. \(\Delta\)R represents RING domain deletion. A) Error bars represent standard deviation (n=4). B) Error bars represent standard deviation (n=3).

4.2.1 Cell cycle analysis

To explore how loss of Msl2 cause the cell cycle progression disturbances, the cell cycle distribution from the above cell lines were analysed by flow cytometry.

First, asynchronous DT40 wild-type, Msl2\(^{-/}\) and Msl2-rescue cells were stained with propidium iodide (PI) for analysis of cellular DNA content by flow cytometry. There is no significant change on cell distribution in the different phases of the cell cycle between each cell line (Figure 4.2A).
Then, we looked in more detail at the cell cycle of these cells. The cells were stained with H3S10ph antibody for flow cytometry analysis to investigate the mitotic cell population. Interestingly, we observed an increase in the mitotic index in the Msl2\(^{-}\) cells (5\%) compared to wild-type (4\%), and this was reduced to wild-type levels in the Msl2-rescue cell line (Figure 4.2B and C). This increase may partially explain the defect in proliferation, and may indicate an underlying problem in the execution of mitosis or the preceding S-phase.
4.3 Histone modifications perturbed in Msl2 knockout cells

As mentioned above, human MSL2 is a component of the hMSL complex with hMOF (Mendjan et al., 2006; Smith et al., 2005), and depletion of hMOF causes a number of phenotypes including a G2/M arrest reminiscent of the delay we observed (Figure 4.2), as well as DNA repair defects (Li et al., 2010; Smith et al., 2005; Taipale et al., 2005). The Becker lab has recently shown that Drosophila MSL2 ubiquitylates MOF, and other MSL proteins, controlling the stoichiometry of the complex (Villa et al., 2012). We therefore wanted to determine whether chicken Mof was affected by loss of Msl2, and whether these knockout cells had similar defects.

In the absence of a functional antibody to chicken Mof protein, we looked at the acetylation levels of one of its substrates; lysine 16 on histone H4 (H4K16ac), for which it is responsible (Akhtar and Becker, 2000; Smith et al., 2005; Taipale et al., 2005). We prepared nuclear extracts and quantified the level of H4K16ac in wild-type or Msl2 knockout cells by immuno-blotting analysis. We found that the level of this modification in Msl2-/- cells was reduced to ~40% of that in wild-type cells (Figure 4.3A and B). This decrease in H4K16 acetylation returned to normal levels in the Msl2-rescue cell line. This suggests that Msl2 regulates the activity of the Mof enzyme, and/or the stability of the complex, and/or the targeting of Mof to its substrate.

In addition, we looked at several other histone modifications; di-methylation of lysine 79 on histone H3 (H3K79me2), di- and tri-methylation of lysine 20 on histone H4 (H4K20me2/3), di-methylation of lysine 9 on histone H3 (H3K9me2), and phosphorylation of serine 10 on histone H3 (H3S10ph). Interestingly, we found that most of these modifications were reduced, with the exception of H3S10ph, which increased (Figure 4.3A and B), in keeping with the mitotic index data (Figure 4.2C).

To confirm that these changes were caused by loss of Msl2 we investigated the levels of H4K20me2 and H4K20me3 in the Msl2-rescue cell line. We found that
re-expression of Msl2 restored and increased the level of H4K20me2, and partially rescued the decrease in H4K20me3 (Figure 4.3A and B).

As the MOF-MSL complex is involved in transcriptional regulation in both Drosophila and human cells (Rea et al., 2007), we looked at whether these changes were due to differential transcription of the enzymes thought to be responsible for the modifications. This does not appear to be the case, as no significant difference in the levels of transcript was detected for those tested (Mof, Suv420 and Dot1l) according to Q-PCR (Figure 4.3C).

Figure 4.3 Histone modifications are perturbed by loss of Msl2. A) Representative immunoblot analysis of nuclear extracts prepared from wild-type, Msl2−/−, and Msl2-rescue cell lines. Panels were probed with the antibodies indicated. B) Quantification of A). Mean expression levels of the various modifications in the cell lines were quantified and expressed relative to those in wild-type cells after normalization to H3 levels. Error bars represent standard deviation (n≥3). C) Q-PCR showing expression levels of the genes indicated. Expression levels in the cell lines are expressed relative to those in wild-type cells, following normalization to Gapdh.

Given that hMSL1/2 or hMOF regulates HOXA9 and MEIS1 gene transcription (Dou et al., 2005; Milne et al., 2002; Wu et al., 2011), we also tried to measure Hoxa9 and Meis1 gene expression in DT40 cells. Two separate pairs of
primers were used for the amplification of Hoxa9, but unfortunately neither of them worked in our study (data not shown). However, Q-PCR analysis of the Meis1 transcript showed that its expression was not affected by loss of Msl2 (data not shown), which is inconsistent with the observations in hMSL2-depleted human cells. Depletion of hMSL2 protein leads significantly down-regulation of Meis1 expression in HeLa cells (Wu et al., 2011).

4.4 Msl2 involved in DNA damage response

There are numerous observations that implicate the human MSL complex in the DNA damage response: hMOF participates in the autophosphorylation of ATM following IR-induced DNA damage (Sykes et al., 2006). hMOF mediated H4K16ac is required for recruitment of the DNA damage mediator protein MDC1 (Kruse and Gu, 2009). hMOF is involved in both non-homologous end joining (NHEJ) and homologous recombination (HR) DSB repair pathways, affecting the IR-induced foci formation of γ-H2AX, Rad51, MDC1, 53BP1 and hSSB (Sharma et al., 2010). Although there is no direct evidence showing hMSL1’s function in the DNA damage response, its mouse homolog is required for the recruitment of MDC1 to DSB foci (Li et al., 2010). Recently, hMSL2 was found to co-purify with DNA damage transducer DNA-PKcs as part of a human hMOF complex (Sharma et al., 2010).

In our initial characterization of the Msl2 knockout cells, we found that several histone modifications were affected by loss of Msl2 (Figure 4.3). Among them, H4K16ac, H3K79me2 and H4K20me2 have been implicated in the DNA damage response (Botuyan et al., 2006; Greeson et al., 2008; Krishnan et al., 2011; Rea et al., 2007; Wakeman et al., 2012).

To address this, my colleague Simona Moravcová carried out an in vivo end-joining reporter assay to test the involvement of Msl2 in the DDR using DT40 wild-type and Msl2\(^{-}\)\(^{-}\) cells. The data showed that the Msl2\(^{-}\)\(^{-}\) cells have defects in re-ligation of the break ends, this was also demonstrated in mammalian U2OS and GC92 cells (Lai et al., 2013). We therefore decided to
perform clonogenic survival assay to better understand the role of Msl2 in response to DNA damage.

### 4.4.1 Msl2 is required for HR repair

DT40 wild-type, \(Msl2^{+/}\) #2 and \(Msl2\)-rescue cells were treated with increasing doses of IR, olaparib, methylmethane sulfonate (MMS) or ICRF-193 and then grown in methylcellulose media for colony counting (see Section 2.2.3.5). IR causes complex DNA damages, in which, double strand breaks (DSBs) is the most lethal one. They can lead to genomic instability and cell death (Su et al., 2010). MMS, as a methylating agent, is believed to induce replication forks stalling in cells (Groth et al., 2010). Olaparib, known as a specific inhibitor of the Poly(ADP-ribose) polymerase PARP1/2, is selective for HR-deficient cells (Birkelbach et al., 2013; Shaw and Hall, 2013). ICRF-193, as a DNA Topoisomerase II (Topo II) catalytic inhibitor, induces DNA damage repaired exclusively by NHEJ (Iijima et al., 2010).

The sensitivity of \(Msl2\)-deficient cells to these DNA damaging agents is shown in Figure 4.4. \(Msl2\)-null cells did not show any survival defect following IR. And unexpectedly, these cells were not sensitive to ICRF-193 treatment, although they showed the impaired end-joining ability from an \textit{in vivo} end-joining assay. This is probably due to the difference of detection times following DNA damage. For an \textit{in vivo} end-joining assay, the cells were collected just after 24 hours of DNA damage. But for a survival assay, cells are incubated for 7-10 days to facilitate colony formation. These observations suggest that Msl2 functions in the early events of, but is not essential for NHEJ. However, the \(Msl2\)-deficient cells showed a moderate sensitivity to MMS, which was also unexpectedly observed in the rescue line. The reason for this is not known. In contrast, when cells were treated with olaparib, an impaired repair capability was observed in \(Msl2^{+/}\) cells, but this was recovered in \(Msl2\)-rescue cells. According to these survival data, we suggest that Msl2 is probably involved in HR repair.
Chapter 4 Functional analysis of $Msl2^{-/-}$ and $Msl2^{-/-}$-rescue DT40 cells

Figure 4.4 Survival curves of DT40 cells in methylcellulose colony assay. DT40 wild-type, $Msl2^{-/-}$ #2 and $Msl2^{-/-}$-rescue cells were treated with increasing doses of MMS, IR, olaparib or ICRF-193 and growing in methylcellulose media to form colonies. The percentage survival at each dose was normalized to the untreated control. The experiment was repeated in triplicate and the mean and standard deviation are also represented in the survival curves shown above.

4.4.2 Msl2 accumulates in response to DNA damage

4.4.2.1 Msl2 is accumulated by MMS treatment

Given the above, the work in the lab implicated that Msl2 does participate in the DNA damage response, so we questioned how Msl2 behaves in response to DNA damage. In DT40 cells, we could not monitor endogenous Msl2 so used the Msl2-rescue cell line instead.

First, DT40 wild-type, $Msl2^{-/-}$ and $Msl2^{-/-}$-rescue cells were treated with 25 μg/ml MMS, and harvested at the following time points: 0, 1, 3 and 6 hours. Whole cell extracts from these time points were analysed by western blotting. FLAG-Msl2 was detected by anti-FLAG antibody. The phosphorylation of H2AX on serine 139 ($\gamma$H2AX) was used to indicate the DNA damage response over this timecourse. As shown in Figure 4.5A, Msl2 protein levels significantly increase after 1 hour of MMS treatment and are maintained up to 6 hours. Interestingly, following DNA damage induced by MMS, $Msl2^{-/-}$-rescue cells display similar levels of $\gamma$H2AX to wild-type cells. But both Msl2 knockout cell
lines had higher levels of γH2AX compared to wild-type cells. This suggests that cells lacking Msl2 suffer greater DNA damage from MMS.

Next, we were interested in the kinetics of Msl2 protein accumulation after MMS treatment. *Msl2*-rescue cells were treated with 50 μg/ml MMS for 3 hours, after which the drug was washed out and cells cultured for up to 24 hour. The cells from the indicated time points were collected, whole cell extracts were prepared and analysed by western blotting. We observe that by 3 hours of MMS treatment (and possibly before) HA2F-Msl2 protein has accumulated compared to the undamaged cells. Protein levels peak between 6 and 12 hours, and have begun to drop off by 24 hours, but still not to back to pre-damage condition (Figure 4.5B). So, we concluded here that Msl2’s expression or protein stability is affected by MMS-induced DNA damage. However, the mechanism behind this is not clear.

Figure 4.5 Msl2 is accumulated following MMS-induced DNA damage. A) Immunoblot analysis of DT40 wild-type, *Msl2*<sup>−/−</sup> and *Msl2*-rescue cells with the antibodies indicated, before and after 25 μg/ml MMS at the times indicated. B) Immunoblot analysis of DT40 *Msl2*-rescue cells with the antibodies indicated. Cells were treated with 50 μg/ml MMS for 3 hours followed by wash out. Cells were harvested at the times indicated. Arrowhead indicates HA2F-Msl2, the upper band is a non-specific anti-FLAG artifact seen in DT40 whole cell extract.
4.4.2.2 Msl2 is stabilised by IR treatment

In section 4.4.2.1, we found that the amount of Msl2 protein is increased in response to MMS-induced damage. So the next question for us is whether Msl2 is specifically involved in MMS-induced DNA damage response or whether it has a more general role in the DNA damage response.

To address this, we also exposed these cells to 5 Gy of ionizing γ-irradiation (IR) and follow the protein level of Msl2 over a timecourse of 12 hours (Figure 4.6A). The level of γH2AX was used to indicate the DNA damage response over this timecourse. We observed that γH2AX levels began to increase after 30 minutes, peaked after 3 hours, and returned to pre-damage levels by 9 hours.

Interestingly, similar to what we found in cells treated with MMS, HA2F-Msl2 protein appeared to accumulate under these conditions; levels began to increase after 1 hour and peaked at 3 hours, before returning to pre-damage levels after 9 hours. This increase was due to some post-transcriptional effect as there was no significant change in Msl2 mRNA levels over the time course (Figure 4.6B).

To investigate this stabilisation further, we treated the DT40 Msl2-rescue cells with the proteasome inhibitors MG132 and ALLN (N-Acetyl-L-leucyl-L-leucyl-L-norleucinal). Both treatments caused an increase in abundance in the amount of HA2F-Msl2 protein as detected by anti-FLAG antibody (Figure 4.6C). Intriguingly, alongside the accumulation of HA2F-Msl2, we noticed the appearance of a band slightly larger than the endogenous protein, suggestive of some post-translational modification.
Chapter 4 Functional analysis of Msl2<sup>−/−</sup> and Msl2<sup>−/−</sup>-rescue DT40 cells

Figure 4.6 Msl2 is stabilized following DNA damage. A) Immunoblot analysis of DT40 wild-type and Msl2-rescue cells with the antibodies indicated, before and after 5 Gy IR at the times indicated. Arrowhead indicates HA2F-Msl2, the upper band is a non-specific anti-Flag artifact seen in DT40 whole cell extract. B) Q-PCR analysis of the Msl2-rescue cell line showing HA2F-Msl2 expression after 5 Gy IR treatment at the times indicated. Error bars represent standard deviation (n=3). C) Immunoblot analysis of the Msl2-rescue cells following treatment with DMSO (vehicle), 50 or 100 µM ALLN or 3 µM MG132 for 8 hours. HA2F-Msl2 is indicated with an arrowhead. Modified HA2F-Msl2 is indicated with an open arrowhead.

Taken together, this suggests that Msl2 is stabilized in response to MMS- or IR-induced DNA damage. We questioned whether Msl2’s E3 ligase activity is required for the stabilisation of Msl2. To address this, the two mutant rescue cell lines were treated with 5 Gy IR over a timecourse of 6 hours. Compared to the wild-type Msl2 (Figure 4.7A), the Msl2 (C44R) protein displayed the similar accumulation pattern following IR: the level increases from 1 hour, lasts for 2 hours, then returns to the pre-damage level by 6 hours. The Msl2 (ΔR) protein also gets stabilised after 1 hour, but the accumulation persists and protein levels remain elevated after 6 hours (Figure 4.7B and C). This is probably due to the mislocalization of the Msl2 (ΔR) protein. Unpublished data from the lab show that this mutant protein exists almost equally in both nuclear and cytoplasmic fractions. The cytoplasmic Msl2 (ΔR) may affect the kinetics of total Msl2
protein, as it cannot be properly regulated as its nuclear fraction in response to DNA damage.

Figure 4.7 Both Mutant Msl2 proteins are stabilized following DNA damage. Immunoblot analysis of the whole cell lysates from Msl2-rescue (A), Msl2-ΔR (B) and Msl2-C44R (C) cells with the antibodies indicated, before and after 5 Gy IR at the times indicated.

4.5 Discussion

Human MSL2 is a component of the human MSL complex, along with hMOF, hMSL1, hMSL3, and NUP153 (Mendjan et al., 2006; Smith et al., 2005). To date, the direct studies on hMSL2 are few. Since hMSL2 was first identified as an E3 ligase, only two substrates, p53 and histone H2B, have been reported (Kruse and Gu, 2009; Wu et al., 2011). To better investigate the biological function of MSL2, chicken Msl2 knockout and rescue cells were generated and used in this study.

The Msl2 knockout cells appear morphologically normal, and a proliferation assay showed that these cells suffer a slight growth defect (Figure 4.1A). Normal proliferation is restored in the Msl2-rescue rescue cells, but not in the mutant Msl2 rescue cells (Figure 4.1B), suggesting that Msl2’s contribution to normal DT40 cell proliferation relies on its E3 ligase activity. Flow cytometry analysis demonstrated that, compared to wild-type cells, cells null for Msl2 accumulate more in mitosis.

Both hMOF and hMSL2 are involved in histone modification: hMOF is responsible for H4K16ac (Gupta et al., 2008; Li et al., 2010; Smith et al., 2005; Taipale et al., 2005); and hMSL2 is reported to ubiquitylate histone H2B K34 (Wu et al., 2011). Given these findings, we investigated a selection of histone modifications in our DT40 cell lines.
In human cells, both hMOF and hMSL1 depletion cause the down-regulation of H4K16ac (Smith et al., 2005). Here we show that elimination of chicken Msl2 also leads to the reduction of this modification (Figure 4.3). This reduction in H4K16ac may be due to lack of MSL complex formation/stability in the absence of MSL2 as previously described in Drosophila (Copps et al., 1998; Straub et al., 2013), or, as hMSL2 possibly ubiquitylates hMOF (Lai et al., 2013), as was recently shown in Drosophila (Villa et al., 2012), it is conceivable that this hypothetical ubiquitylation promotes hMOF’s activity towards H4K16.

In human cells, H3K4me3 and H3K79me2 are also reduced in the hMSL2 knockdown cells. It is not due to the transcription regulation of MLL or DOT1L genes (Wu et al., 2011). In our studies, we did not check H3K4me3, but the change of H3K79me2 in the Msl2 knockout cells is consistent with the observation in human cells (Figure 4.3). The reduction in H3K79me2 is probably due to presumed loss of hMSL2 mediated H2BK34ub and the subsequent loss in stimulation of the DOT1L methyltransferase as previously reported (Wu et al., 2011).

In addition to H4K16ac and H3K79me2, several important histone modifications are also disrupted; H4K20me2 and H4K20me3, with levels reduced to ~25-50% that in wild-type cells (Figure 4.3B). This is surprising in light of our finding that the cells only show slight growth defect. However, none of these modifications were completely lost and the residual levels may be sufficient for the cells to grow almost normally.

The Msl2 knockout cells then were found deficient in the NHEJ ligation repair, which strongly supports the proposed role of Msl2 in the DNA damage response. In our lab, using in vivo end-joining reporter assay, Simona Moravcová observed that the DT40 cells lacking Msl2 had an impaired ability (25-50%) to religate the digested DNA compared to wild-type cells. Again, using the same assay, she found U2OS cells depleted of hMSL2 had an impaired ability (72%) to repair the digested DNA compared to control siRNA treated cells (Lai et al., 2013). This was further proven in an established intrachromosomal NHEJ
substrate-based system (Rass et al., 2009) in collaboration with Dr. Yvan Canitrot (Lai et al., 2013). Depletion of hMSL2 from this system resulted in a reduction (58%) in the frequency of end-joining. Given the above, both chicken Msl2 and human MSL2 contribute to end-joining repair. The requirement of Msl2 for a fully functional NHEJ repair pathway is complex but probably partly converges at the recruitment of 53BP1, a key protein in the NHEJ pathway (Bunting et al., 2010; FitzGerald et al., 2009; Noon and Goodarzi, 2011). 53BP1 is recruited to DNA damage sites by different means: acetylation of H4K16 (Li et al., 2010); di-methylation of H3K79 (Huyen et al., 2004; Wakeman et al., 2012); and di-methylation of H4K20 (Greeson et al., 2008; Pei et al., 2011; Wakeman et al., 2012) have all been shown to (directly or indirectly) recruit 53BP1. We have observed reduced levels of all these modifications in the Msl2 knockout cells (Figure 4.3). In addition to histone modifications, 53BP1 also gets recruited to damage sites by interaction with the mediator protein MDC1 (Eliezer et al., 2009; Mok and Henderson, 2012). Others have previously shown that depletion of hMOF or hMSL1 causes a loss of recruitment of MDC1 to damage foci (Li et al., 2010). Our data further implicates that knockdown of hMOF or hMSL1 significantly decreases the MDC1 expression at protein level, but not at transcription level (see Appendix 1). This explains the observation that hMSL1 and hMOF are required for the efficient recruitment of MDC1 into foci after DNA damage (Li et al., 2010); without these proteins MDC1 is probably degraded. It is interesting that both hMSL1 and hMOF appear to get modified in the presence of hMSL2 in response to DNA damage (Lai et al., 2013). One could speculate that this modification, possibly ubiquitination, may regulate these proteins’ interaction with/activity towards MDC1 and somehow promote its recruitment or stabilization at damage sites. Whereas, we hypothesize that stabilized Msl2/hMSL2 could facilitate the histone modifications described above, promoting the recruitment or accumulation of 53BP1 leading to NHEJ-mediated repair.
As we mentioned in Chapter 1, the NHEJ and HR pathways are the two main repair pathways for DSB repair (Ciccia and Elledge, 2010). In mammalian cells, NHEJ is thought to be the major repair pathway throughout most of the cell cycle, repairing more than 90% of DSBs, whereas, HR is generally restricted to late S and G2 phases of the cell cycle (Kanaar et al., 2008a; Rothkamm et al., 2003). To better understand the role of Msl2 in DNA damage, we performed the clonogenic survival assay to evaluate the repair capacity of Msl2 to IR, MMS, ICRF-193 or olaparib. The cells without Msl2 do not show any defects to repair IR- MMS- or ICRF-193-induced DNA damage. However, they are sensitive to the PARP1 inhibitor olaparib, suggesting that Msl2 is involved in HR repair. Given the above, we wanted to study how Msl2 itself behaves in response to damage. In the absence of a functional antibody, the Flag-Msl2 served as surrogate for the endogenous Msl2 protein. To our surprise, Msl2 is stabilized in response to IR and MMS-induced damage (Figure 4.6 and 4.7). As ALLN/MG132 treatment also leads to Msl2 accumulation, this stabilization appears to be dependent on the avoidance of some proteolytic action against Msl2 that may be constantly keeping Msl2 at low basal levels. In keeping with the DT40 result is the accumulation of hMSL2 protein when cells are treated with IR or ALLN (Lai et al., 2013). As the kinetics of Msl2 accumulation closely correlates with γH2AX levels, these results suggest that soon after, or coincident with the formation of γH2AX foci, Msl2/hMSL2 accumulates in the cell. Based on our Q-PCR, this accumulation is speculated due to some post-translational modification. MDM2, another E3 ligase, has been found to be phosphorylated near the RING domain by ATM when DNA damage occurs. This phosphorylation prevents MDM2 polyubiquitinating its specific substrate, p53, promoting the stabilization of p53 in response to DNA damage (Cheng et al., 2011). Together with our western blot result, we suggest that Msl2, like MDM2, is probably phosphorylated rather than ubiquitinated, as the shift of modified Msl2 is quite small.
This modification possibly mediates the inhibition of proteolysis of Msl2/hMSL2 by the proteasome. As MSL2 in *Drosophila* has been shown to polyubiquitylate itself, MSL1, MSL3 and MOF to regulate the MSL complex stoichiometry via proteasome-mediated degradation (Villa et al., 2012), this stabilization of Msl2 may extend to other complex members. In addition, the nuclear localisation of Msl2 is revealed by western blot using the *Msl2*-rescue cells (Figure 3.11), which is the same as its human homolog. This subcellular distribution facilitates a rapid response of Msl2/hMSL2 to DNA damage. Further, both mutant Msl2 proteins displayed the similar accumulation pattern following IR (Figure 4.7B and C), suggesting the enzymatic activity of Msl2 is not required for its stabilisation in DNA damage response and also that the potential posttranslational modification sites are outside the RING domain region, so the modified residue(s) is not in the N-terminal 107 amino acids, further supporting our hypothesis that Msl2 probably takes the similar mechanism as MDM2 to respond to DNA damage in the early stage of the DDR.
Chapter 5 Conclusions and future perspectives

In this study, we have cloned the chicken Msl2 cDNA. We also mapped the genomic locus of Msl2, which facilitated a cloning strategy that allowed the generation and characterisation of DT40 Msl2 knockout cell lines. Based on the data presented in this thesis, we have determined that Msl2 is a novel component in the vertebrate DNA damage response pathway. We found that Msl2 is stabilized early in response to IR- or MMS-induced DNA damage. The kinetics of this stabilization closely follow γH2AX appearance. Both IR and MMS are well known to cause the double-strand breaks in cells and ATM is the main sensor protein to initialize the response to the damage (Canman et al., 1998; Carrozza et al., 2009). Interestingly, we also found that the Msl2 knockout cells are sensitive to the PARP1 inhibitor olaparib, indicating Msl2 is required for HR repair. The inhibition of PARP has been shown to induce DNA double strand breaks, which are resolved in an ATM-dependent pathway (Aguilar-Quesada et al., 2007). The above information implies that Msl2 probably functions as a mediator in DSB DNA damage response, functioning downstream of the activated ATM. Indeed Msl2 itself may be phosphorylated by ATM as we have observed a migration shift of Msl2 in response to damage that would correlate with phosphorylation. The potential role of Msl2 in response to DNA damage is schematically represented in Figure 5.1.

In Drosophila, in order to sustain the MSL complex stoichiometry, MSL2 polyubiquitylates itself and other MSL subunits, mediating their proteasome degradation (Villa et al., 2012). In keeping with this, we found that the loss of Msl2 perturbs the H4K16 acetylation mediated by Mof; while overexpression of Msl2 also affects the same modification. We suggest that the expression level of Msl2 is tightly regulated, and may be kept at a low basal level sufficient to maintain normal cellular processes.
Following on from these conclusions, we can propose a model for Msl2’s role in the DDR: under DNA damage conditions, ATM, sensing the damage, autophosphorylates itself and activates DNA damage signalling transduction. The activated ATM may phosphorylate Msl2 leading to its stabilization. This phosphorylation also possibly regulates Msl2’s (poly)ubiquitylation activity and/or interaction with other Msl proteins promoting their stability. The stabilized Msl2 (and possibly a stabilized Msl complex) is recruited to the vicinity of double-strand breaks via its C-terminal CXC DNA binding domain or by some unknown interacting proteins. Binding of the Msl complex to the damaged chromatin mediates different events to regulate the DNA damage response:

Mof promotes the acetylation of H4K16, which generates an open conformation of chromatin to expose H2B K34 for monoubiquitination by Msl2/Msl1. These two histone modifications may mediate DNA repair in several ways. Acetylation can reduce the affinity of histones for DNA making it more accessible for the direct recruitment of downstream DNA repair machinery, or possibly expose other ‘damage’ modifications such as H3K79me. H4K16ac and H2BK34ub may also regulate other histone modifications via trans-tail crosstalk, such as H3K79me2 and H4K20me2/3, to indirectly recruit downstream DNA repair machinery. H3K79me2 and H4K20me2/3 have been shown to recruit 53BP1 to DNA damage sites.

Activated ATM phosphorylates H2AX (γH2AX) to recruit MDC1 in response to DNA damage. 53BP1 gets recruited to damage sites in an H2AX- and MDC1-dependent manner. The reduced recruitment of MDC1 has been reported in MOF- or MSL1-depleted cells. Based on the data shown in the Appendix, the less MDC1 foci formation reported as a recruitment defect is incorrect and is probably due to the down-regulation of MDC1 at the protein level in MOF- or MSL1-depleted cells. This suggests that MOF and MSL1 are required for the stability of MDC1 via post-translational regulation. As MOF and MSL1 are
regulated by MSL2, we hypothesize that MsI2 is also involved in DNA damage response by indirect regulation of MDC1 and downstream factors.

![Image](image.png)

**Figure 5.1 Schematic representation of the potential function of Msl2 suggested in DNA damage response.** A) In physiological conditions, MSL2’s autoubiquitylation and its polyubiquitination of other MSL subunits maintain the MSL complex stoichiometry. B) Following DSB damage, phosphorylated MSL2 stabilizes and mediates the recruitment of the MSL complex to DNA lesions, initializing a series of cellular processes in response to DNA damage.

To verify the role of Msl2 in our proposed model, we need to address whether Msl2 acts alone or together with the other components of the Msl complex, or with some unknown interacting proteins, to facilitate the DDR. So we next plan to combine co-immunoprecipitation and mass spectrometry approaches to identify new Msl2 interacting or substrate proteins using Msl2 knockout and rescue cell lines treated with or without IR. The potential interactors identified would shed light on the functional role of Msl2 in DNA damage response.

The MSL2 protein has been identified as an E3 ubiquitin ligase in human and Drosophila, of which the characterised substrates include p53, histone H2BK34 and the MSL members; MOF, MSL1 and MSL3. In the lab, a blast search has been done in human genome database, using the sequence surrounding H2BK34. Interestingly, from the first seven relevant (containing the lysine corresponding...
to K34) protein hits, two of them are located in 53BP1, indicating 53BP1 is possibly a substrate of hMSL2 (Lai et al., 2013). It is well known that 53BP1 plays an important role in DNA damage response. If it is identified as a substrate of Msl2, which will greatly help us to determine the position of Msl2 in DNA damage response. Based on these, we want to identify the E3 ligase activity of chicken Msl2 on hMOF and 53BP1.

Taken together, the studies on the enzymatic activity of Msl2 and the potential interacting proteins of Msl2 in the future will help us to better understand the mechanism of Msl2 in DNA damage response.
References


Bekker-Jensen, S., Danielsen, J. R., Fugger, K., Gromova, I., Nerstedt, A., Bartek, J., Lukas, J.,


phosphorylation of BRCA1 in the DNA damage response to double-strand breaks. Science 286, 1162-1166.


References

References


References

Biochim Biophys Acta 1490, 170-174.
References

819-824.
References

844-847.
References


References


Appendix 1 MDC1 downregulated in hMSL1-depleted cells

siRNA mediated depletion of hMSL1 leads to a dramatic reduction in the levels of MDC1 protein, whereas depletion of hMSL3 did not have any effect. Furthermore, due to no significant change of MDC1 expression at transcript level, we suggest that hMSL1 is required for the post-translational regulation of MDC1 stability.

Figure A1.1 MDC1 downregulated in hMSL1-depleted cells. U2OS cells were transiently transfected with siRNA against hMSL1 or hMSL3 for 48 h, control-siRNA included as a negative control. mRNAs from indicated cells were used for Q-PCR analysis. Whole cell lysates from indicated cells were used for western blot analysis. A) Q-PCR analysis of expression of hMSL1 and hMSL3. B) Western blot analysis of MDC1 downregulation. C) Q-PCR analysis of expression of hMSL1, hMSL2 and MDC1.
Appendix 2 Detection of chicken Msl2 and Mof expression by human antibodies

A2.1 Introduction

Currently, no commercial antibody against chicken Msl2 is available. In the lab, there are three commercial polyclonal and one homemade monoclonal anti-hMSL2 antibodies. One of the commercial antibodies is a rabbit anti-hMSL2 from Abnova. The other two are mouse anti-hMSL2 and rabbit anti-hMSL2 from Abcam. As none of the commercial antibodies recognised hMSL2 by western blot, we decided to make our own monoclonal antibody. A bacterial derived hMSL2 fragment comprising amino acid residues from 86 to 412 (the sequence between RING and CXC domains) was used as immunogen. The supernatants from 26 different clones were obtained from Dundee Cell Products (UK) and used for detection of the endogenous or exogenous hMSL2 protein by Simona Moravcová. Five of the supernatants could recognise both endogenous and exogenous hMSL2 with relatively little background, with the supernatants from clones 8A4, 4F12 and 8D2 giving the best results. The supernatant from 4F12 was used in all further experiments to detect hMSL2.

We compared amino acid sequence of chicken Msl2 with this antigen region of hMSL2 to check the degree of conservation between them. The target sequence is less conserved than the N-terminal RING finger domain or C-terminal CXC domain, but chicken Msl2 still shows 77% identity and 87% similarity with the human sequence. With such high homology, we hoped that the supernatants generated from this antigen would also work for chicken Msl2.

As mentioned in chapter 3, the chicken Msl2 gene was previously successfully disrupted in DT40 cells and two knockout cell lines were identified by Southern blotting and Q-PCR. These lines provide a useful tool for testing whether the anti-hMSL2 antibody recognises the chicken homologue.
A2.2 hMSL2 antibody test in DT40 cells

The whole cell lysates obtained from DT40 wild-type, $Msl2^{+/}$ #1 and $Msl2^{-/-}$ #2 cells were used for our initial antibody test. Immunoblot analysis revealed that the human commercial antibodies recognised neither hMSL2 protein (data not shown) nor chicken Msl2 protein (Figure A2.1).

Figure A2.1 Testing of commercial anti-hMSL2 antibodies in DT40 cells. Western blot analysis of endogenous Msl2 in DT40 wild type and $Msl2^{-/-}$ #1 cells with three different polyclonal anti-hMSL2 antibodies from 2 different companies. U2OS cell were used as a control. The predicted size of chicken Msl2 is ~75 kD. A) anti-hMSL2 mouse polyclonal antibody (Abcam) used. B) anti-hMSL2 rabbit polyclonal antibody (Abcam) used. C) anti-hMSL2 rabbit polyclonal antibody (Abnova) used.

Three monoclonal hybridoma supernatants were then tested for their ability to detect endogenous chicken Msl2 protein using DT40 whole cell lysates. No positive signal was observed following western blotting for any of the supernatants (data not shown). In Drosophila, MSL-2, together with MSL-1, is involved in the ubiquitination of other MSL proteins controlling their abundance and regulating the stoichiometry of the MSL complex. Its expression therefore is tightly regulated (Villa et al., 2012). We suspected that Msl2 expression displays similar low abundancy and is also tightly regulated in vertebrate cells, as the endogenous protein is not detected in whole cell lysates. To address this, I used the DT40 nuclear extracts for further immunoblot analysis (Figure. A2.2). A similar band pattern was seen on each film, the size of Msl2 is predicted around 75 kD, however above which, only a strong 90-kD band occurred in all cell samples, suggesting the homemade monoclonal anti- hMSL2 antibody failed to
recognize endogenous chicken Msl2.

![Figure A2.2 Testing of monoclonal anti-hMSL2 antibody in DT40 cells.](image)

Figure A2.2 Testing of monoclonal anti-hMSL2 antibody in DT40 cells. Nuclear extracts from DT40 wild type, Msl2<sup>−/−</sup> cells were used for western blot analysis of endogenous Msl2. Three hybridoma supernatants: 4F12 (A), 8A4 (B) and 8D2 (C) were tested. The predicted size of chicken Msl2 is around 75 kD.

**A2.3 hMSL2 antibody test of exogenous chicken Msl2 in U2OS cells**

To exclude the possibility that human antibodies did not recognise the endogenous chicken Msl2 due to a low abundance limitation, I transiently transfected the pCDNA3.1-HA2F-Msl2 plasmid into human U2OS cells in order to over-express the HA2F-Msl2 fusion protein for an antibody test. The predicted size of the fusion protein is ~75 kD also.

The whole cell lysates from U2OS cells transiently transfected with empty flag plasmid (pCDNA3.1-HA2F), flag-tagged Msl2 (pCDNA3.1-HA2F-Msl2) and flag-tagged hMSL2 (pCDNA3.1-HA2F-hMSL2) were used for western blotting analysis, with flag-tagged hMOF (pCDNA3.1-HA2F-hMOF) expression used as a positive control. As shown in Figure A2.3D, the exogenous protein expression was clearly detected with anti-FLAG antibody.
Figure A2.3 Testing of monoclonal anti-hMSL2 antibody in Msl2 transfected U2OS cells. U2OS cells were transiently transfected with pCDNA3.1-HA2F-Msl2 (wt), pCDNA3.1-HA2F-hMSL2, pCDNA3.1-HA2F-hMOF, or pCDNA3.1-HA2F vector. 48h after transfection, whole cell lysates from cells were used for western blot analysis using the mouse anti-FLAG antibody or three different supernatants against hMSL2 (8D2, 8A4, 4F12). The apparent size of exogenous chicken Msl2 (wt) is ~75 kD, ~80 kD for exogenous hMSL2, 54 kD for exogenous hMOF and 75 kD for endogenous hMSL2. HA2F-hMSL2 is indicated with an open arrowhead. Endogenous hMSL2 is indicated with an arrowhead. HA2F-Msl2 is indicated with an asterisk.

To confirm the above observations, we repeated the experiment using the supernatant 8D2 and included the ponceau S staining data to confirm the equal loading of protein samples (Figure A2.4). Again, all the exogenous proteins were successfully expressed in U2OS cells. However, there was no significant signal difference around 75 kD between different cell sample, suggesting that the homemade monoclonal antibody does not recognize the chicken protein.
Figure A2.4 Testing of the anti-hMSL2 antibody 8D2 in Msl2 transfected U2OS cells. U2OS cells were transiently transfected with pCDNA3.1-HA2F-Msl2 (wt), pCDNA3.1-HA2F-hMSL2, pCDNA3.1-HA2F-hMOF, or pCDNA3.1-HA2F vector. For hMOF transfection, 0.1 and 1.0 μg of target plasmid were used separately. 48h after transfection, whole cell lysates from cells were subjected to western blot analysis. Blots were probed with monoclonal anti-hMSL2 clone 8D2 (upper panel) anti-FLAG antibodies (middle panel) or stained with ponceau-S. The apparent size of exogenous chicken Msl2 (wt) is ~75 kD, ~80 kD for exogenous hMSL2, 54 kD for exogenous hMOF and 75 kD for endogenous hMSL2. HA2F-hMSL2 is indicated with an open arrowhead. Endogenous hMSL2 is indicated with an arrowhead. HA2F-Msl2 is indicated with a star.

In summary, none of the hMSL2 antibodies available in the lab could be used in direct recognition of chicken Msl2 at the protein level.

A2.4 hMOF antibody test of Chicken homolog in DT40 cells

The MSL complex is conserved during evolution. The chicken homologs of MSL1, MSL2 and MSL3 are annotated on the NCBI and Ensembl databases. A homolog of MOF was not present. However, we did find a short fragment of predicted Mof in the NCBI database (Accession: XM_001236077). Using this Mof sequence, we did another two rounds of Blast in the chicken EST database and three overlapping chicken EST sequences were discovered. When they were combined together (see Appendix 3), they encoded a polypeptide with high homology compared to hMOF protein. As shown in Figure A2.5, this assembled
chicken Mof protein shares 86% identity and 90% similarity with hMOF, but lacks 30-40 amino acids on both the N-terminus and the C-terminus.

Figure A2.5 Alignment of the assembled chicken Mof sequence and hMOF.

Currently, as is the situation for chicken Msl2, no commercial antibodies against chicken Mof are available. In the lab, a homemade mouse monoclonal anti-hMOF antibody was already produced by using bacterial recombinant full length GST-tagged hMOF as an immunogen. The supernatants from several hybridoma clones were proven to recognise endogenous and exogenous hMOF protein in human cells. So I tested these supernatants in DT40 cells. The whole cell lysates from DT40 wild-type and Msl2−/− #1 cell line were used, and the whole cell lysate from human U2OS cell line was included as a positive control. As shown in Figure A2.6, for each supernatant, endogenous hMOF was detected as a band of ~52 kD in size. We did not observe any positive band in DT40 wild-type cells. However, a band with the similar size as hMOF was detected in DT40 Msl2−/− cells using anti-hMOF supernatants 8E3 and 4G4.
Based on these data, we questioned if the hMOF antibody has cross reactivity in chicken species.

Figure A2.6 Testing of anti-hMOF antibody supernatants in DT40 cells. Whole cell lysates from DT40 Wild-type and MsI2−/− #1 cells were used for western blot analysis to detect the predicted endogenous chicken Mof using the anti-hMOF supernatants 4G4, 7D1, 7G11, 8B12 and 8E3, including U2OS cells as a positive control. The predicted size of hMOF is ~52 kD. Endogenous hMOF is indicated with an open arrowhead.
Appendix 3 Partial length chicken Mof sequence

1. Partial length chicken Mof cDNA based on Blast researches

<table>
<thead>
<tr>
<th>Start</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TCTTCCCCCC AACATGGCGG ACCCTCCGGC GCTCTCCGCT CCTCCGGGAA CGGGGAAACG</td>
</tr>
<tr>
<td>61</td>
<td>TACCTGTGCC GGAGGCCGGGA CGGCAGCTGG CACTCAGCAG AGGTGATTCA GTCGCGGCTG</td>
</tr>
<tr>
<td>121</td>
<td>AACGAGCAAG AGGGCCGGGA AGGAGTTTCCT ACAGAGTCGC TGGGCTTTAA CCGCCGGCTTG</td>
</tr>
<tr>
<td>181</td>
<td>AACGAGCAGG AGGGCCGCGA GGAGTTCTAC GTGCACTACG TCGGCTTAA CCGCCGGCTTG</td>
</tr>
<tr>
<td>241</td>
<td>GACGAATGGG TGGACCGGAA CCGTTTGGCG CTGAGCAAAAA CGCTGAAGGA CGCCGTGAGC</td>
</tr>
<tr>
<td>301</td>
<td>AAAAACTCGG AGCAGTTTTT GGGGAAACTC CCAGAACAGC CGGAGAGGAA AATCACCCGG</td>
</tr>
<tr>
<td>361</td>
<td>CCCACCAGCG CGCGCCTGGGA GAAGGACCAC GAAGCGACAA GCCAAGTGGA ATACGTTGGAC</td>
</tr>
<tr>
<td>421</td>
<td>AAAATCCACA TCGGTCACCT TGAGATCGAC GCGTGGTATT TCTCCCCCTT CCCCCGAGAT</td>
</tr>
<tr>
<td>481</td>
<td>TGCCAGGAGC GCCCAACGCT GTGGATCTGC GAGTACTGCC TCAAGTCAAC GAAGTACGAG</td>
</tr>
<tr>
<td>541</td>
<td>CGCACCTACC GCCTCCACCT GGGTCAAGTC CGACTCGGG AGGAGAAACTC CCAGAACAGC</td>
</tr>
<tr>
<td>601</td>
<td>TACCGGAGAC GCACAATCTC TGGTATGAG GTGGACGGCA AAGACCACAA GATCTACTGC</td>
</tr>
<tr>
<td>661</td>
<td>CAGAACCTCT CACCTGCCTGG CGACTGGTTT GTGACCGCTG TGTCCTCAGA AATAGCTGGAC</td>
</tr>
<tr>
<td>721</td>
<td>GAGCCCTTCC TTTTCTACCT GCTGACCGGAC GTGGACCGGG AAGAGCCGAC ATGGAAGATG</td>
</tr>
<tr>
<td>781</td>
<td>ACTTCTCCAA GGGAGAGGAC GTCCTCCGAT TGGGAAACAAC TGGGCTCTTC CTTCCCCGAT</td>
</tr>
<tr>
<td>841</td>
<td>CGGCCGATCC AAGCGGCAGA ATACCGGAAG TTTCTCAGTA CCTTCAGTGA TGAGCTGCTG</td>
</tr>
<tr>
<td>901</td>
<td>AAGCTGGAGA GCACCTGGGG GTCCCCCGAA AAGAGCTGGG CGAATTCGG AAAAACTGGG</td>
</tr>
<tr>
<td>961</td>
<td>TGACCGACTCT ACGGAGGACT GGTGGTCTGC GAGATCTGAC GCGAATCCCG CGGACACGCTG</td>
</tr>
<tr>
<td>1021</td>
<td>TCCATCAAGG ACCTAGCCAC CTCAGGCTCC TCAACATGG TGAGATCCTG GAAAGCCAGC</td>
</tr>
<tr>
<td>1081</td>
<td>CACGTGACCT GCCTCCACCT CAAACCTGGG AGGGAGACCC TCAAAAGCCC CCAGTACAAG</td>
</tr>
<tr>
<td>1141</td>
<td>AAGCCCCCACA TACCCGTTAA T</td>
</tr>
</tbody>
</table>
Appendix 4 Scientific communications

Publication

Posters

Msl2 Is a Novel Component of the Vertebrate DNA Damage Response

Zheng Lai1*, Simona Moravcová1*, Yvan Canitrot2, Lukasz P. Andzejewski1, Dervla M. Walshe1, Stephen Rea1*

1 Centre for Chromosome Biology, School of Natural Sciences, National University of Ireland, Galway, University Road, Galway, Ireland, 2 Université de Toulouse, UPS, CNRS, Toulouse, France

Abstract

hMSL2 (male-specific lethal 2, human) is a RING finger protein with ubiquitin ligase activity. Although it has been shown to target histone H2B at lysine 34 and p53 at lysine 351, suggesting roles in transcription regulation and apoptosis, its function in these and other processes remains poorly defined. To further characterize this protein, we have disrupted the Msl2 gene in chicken DT40 cells. Msl2−/− cells are viable, with minor growth defects. Biochemical analysis of the chromatin in these cells revealed aberrations in the levels of several histone modifications involved in DNA damage response pathways. DNA repair assays show that both Msl2−/− chicken cells and hMSL2-depleted human cells have defects in non-homologous end joining (NHEJ) repair. DNA damage assays also demonstrate that both Msl2 and hMSL2 proteins are modified and stabilized shortly after induction of DNA damage. Moreover, hMSL2 mediates modification, presumably ubiquitylation, of a key DNA repair mediator 53BP1 at lysine 1690. Similarly, hMSL1 and hMOF (males absent on the first) are modified in the presence of hMSL2 shortly after DNA damage. These data identify a novel role for Msl2/hMSL2 in the cellular response to DNA damage. The kinetics of its stabilization suggests a function early in the NHEJ repair pathway. Moreover, Msl2 plays a role in maintaining normal histone modification profiles, which may also contribute to the DNA damage response.

Introduction

DNA double strand breaks (DSBs) are a particularly dangerous form of damage, as their inaccurate repair or lack of repair can result in mutations or chromosomal translocations leading to cancer. DSBs can be repaired by either of two processes: non-homologous end joining (NHEJ) or homologous recombination (HR) [1,2]. HR repair occurs in S- and G2-phases of the cell cycle, when it can use the undamaged nearby homologous sister chromatid’s DNA as a template to faithfully repair the break. NHEJ occurs throughout the cell cycle, is faster than HR, and results in ligation of the two broken DNA fragments [1,2]. Both pathways comprise a series of stages that involve a large and growing number of proteins; sensors first detect that there is a double-stranded break in the DNA. Next, mediators and transducers get recruited to damaged chromatin, where they accumulate. The signal is amplified and passed on to effector proteins. These effectors enable cell cycle arrest and the repair of the broken DNA [2,3,4].

The choice of which pathway a cell takes to repair a DSB is dependent on the stage of the cell cycle and the complexity of the damage, and is crucial to the damaged cell. Perturbation in the balance between HR and NHEJ can result in disease, but can also be exploited in the treatment of cancer [1]. One of the proteins regulating this choice is 53BP1 (p53 binding protein 1). It can inhibit DNA resection, and thus HR repair, promoting the NHEJ pathway [5,6,7]. Following DNA damage it gets recruited to and accumulates at chromatin surrounding the damage site through interaction with methylated histone residues (H3K79me2, H4K20me2) via its tandem tudor domains [8,9,10,11], and through interaction with the damage mediator protein MDC1 via a central core region. Once 53BP1 accumulates it is involved in recruitment of other DDR proteins, facilitating accessibility to the chromatin [12], or otherwise promoting repair [6,7]. However, it is still unclear how the enzymes mediating these 53BP1-recruiting modifications are themselves regulated in response to DNA damage.

Msl2 (male-specific lethal 2) was originally identified in the fruitfly, Drosophila melanogaster, in genetic screens for mutants causing male-specific lethality. Such genes were implicated in dosage compensation, a process that ensures equal amounts of X-chromosomal gene expression between males and females with unequal numbers of this sex chromosome (reviewed in [13,14]). Msl2 was subsequently shown to be a pivotal participant in this process. Its expression is achieved only in male flies and is required for the formation of the MSL complex (also known as the dosage compensation complex) [15], and its initial recruitment to the
male X chromosome [16], whence it mediates the 2-fold increase in transcription of X-linked genes [13,14]. It was recently shown that MSL2 can ubiquitylate MSL1, as well as MSL3 and MOF [17,18]. This ubiquitylation can target these proteins for proteasome-mediated degradation to control MSL complex stoichiometry, but is also proposed to regulate their recruitment to specific chromatin domains [18].

Human orthologues of these MSL proteins exist and they are found in an evolutionary conserved human MSL complex, also known as the hMOF (Males absent on the first, human) complex [19,20,21]. Direct studies on hMSL2 are few; one study found that when overexpressed, it is able to mono-ubiquitylate p53 at lysine 351. This targets p53 for export to the cytoplasm where it induces mitochondrial-dependent apoptosis [22,23]. Mutation of this residue has been reported in a cisplatin-resistant ovarian carcinoma cell line [23]. A second study has recently shown that hMSL2 in tandem with hMSL1 is able to ubiquitylate histone H2B on lysine 34, and that this H2BK34ub directly regulates methylation of H3K4 and H3K79 by trans-tail crosstalk to promote transcription [24].

Other components of the human MSL/MOF complex include hMOF, hMSL1, hMSL3, and NUP153 (Nucleoprotein 153) [19,20,21]. hMOF is involved in transcriptional regulation [19,25], is required for embryogenesis [26], is downregulated in several cancers [27], and importantly here, is known to participate in the DNA damage response (DDR) [28,29,30,31]. hMOF has also been proposed to regulate ATM (Ataxia telangiectasia mutated) function following DNA damage [31]. Moreover, it can acetylate p53 in response to high levels of damage promoting p53-dependent transcription of pro-apoptotic genes [32]. Interestingly, hMSL1 is known to influence hMOF’s H4K16 acetylation activity [21,33], and it has been shown to co-immunoprecipitate with the DNA repair mediator, 53BP1 [33].

The aim of this study was to determine the cellular function(s) of MSL2 in higher eukaryotes, and to investigate whether this function involves MSL2’s interaction with other MSL proteins. Knock-out of chicken Msl2 in DT40 cells has revealed an important role of MSL2 in the DNA damage response. We found that Msl2 is required for normal levels of several histone modifications involved in the DDR, including those that recruit 53BP1. Msl2 is also required for full NHEJ efficiency, as is the human orthologue hMSL2. Both human and chicken proteins are rapidly stabilized in response to DNA damage, and hMSL2 mediates the possible ubiquitylation of 53BP1, hMSL1 and hMOF. These data define Msl2/hMSL2 as a novel player in the NHEJ pathway, acting early in the DDR, and upstream of the modifications and proteins that recruit 53BP1.

**Results**

**Msl2 Knockouts are viable with Minor Growth Defects**

To determine the function of Msl2 in vertebrates we targeted the chicken gene, Msl2 for disruption in DT40 cells. Using available database information we found only one Msl2 gene in...
Histone Modifications are Perturbed in Msl2 Knockout Cells

The exact cause of this growth defect and delay is unclear. As mentioned above, human MSL2 is a component of the hMSL complex with hMOF [20,21], and depletion of hMOF causes a number of phenotypes including a G2/M arrest reminiscent of the Msl2 function, we used genomic PCR to generate targeting constructs that would delete the larger second exon encoding 92% of the protein (Figure 1A).

Our first observation was that the Msl2−/− cells were viable, with morphology similar to wild-type (not shown), demonstrating that Msl2 is not an essential gene in DT40 cells. We then examined the proliferative ability of Msl2−/− cells. We found that cells lacking Msl2 proliferate more slowly than wild-type cells (Figure 1D). Wild-type cells proliferated with a doubling time of 4.9 ± 0.1 hours whereas Msl2−/− #1 and #2 took 9.18 and 9.62 hours respectively, a delay of approximately 20%. The rescue cells displayed a recovery of this delay, having a doubling time of 7.95 hours. We observed an increase in the mitotic index in the Msl2−/− cells (5%) compared to wild-type (4%), and this was reduced to wild-type levels in the Msl2-rescue line (Figure 1E). This increase may partially explain the defect in proliferation.

Histone Modifications are Perturbed in Msl2 Knockout Cells

In the absence of a functional antibody to chicken Mop protein, we looked at the acetylation levels of one of its substrates; lysine 16 on histone H4 (H4K16ac) for which it is responsible [21,30,34]. We prepared nuclear extracts and quantified the level of H4K16ac in wild-type or Msl2 knockout cells by immuno-blotting analysis. We found that the level of this modification in Msl2−/− cells was reduced to ~40% that in wild-type (Figure 2A, B). This decrease in H4K16 acetylation returned to normal levels in the Msl2-rescue cell line. This suggests that Msl2 regulates the activity of the Mof enzyme, and/or the stability of the complex.

In addition we looked at several other histone modifications; di- and tri-methylation of lysine 20 on histone H4 (H4K20me2/3), di-methylation of lysine 79 on histone H3 (H3K79me2), phosphorylation of serine 10 on histone H3 (H3S10ph), and di-methylation of lysine 9 on histone H3 (H3K9me2). Interestingly, we found that the levels of H4K20me2, H4K20me3 and H3K79me2 were reduced to 50% or less than that in wild-type cells. In contrast, H3S10ph levels increased significantly (Figure 2A, B), in keeping with the mitotic index data (Figure 1E). H3K9me2 levels were not affected.

To confirm that these changes were caused by loss of Msl2 we investigated the above modifications in the Msl2-rescue cell line. We found that re-expression of Msl2 restored to almost wild-type levels the defects in modifications observed (Figure 2A, B).

As the MOF-MSL complex is involved in transcriptional regulation in both Drosophila and human cells [35], we looked at whether these changes were due to differential transcription of the enzymes thought to be responsible for the modifications. This does not appear to be the case, as no significant difference in the levels of transcript was detected for those tested (Mof, Suv420 and Dot1l) according to Q-PCR (Figure 2C).

Msl2/hMSL2 Plays a Role in NHEJ

Several of the histone modifications that are affected by loss of Msl2 are implicated in the DNA damage response: H4K16ac [35,36]; H3K79me2 [11] and H4K20me2 [10,37]. hMSL2 was shown to co-purify with the NHEJ repair protein DNA-PKcs as part of a human hMOF complex [29]. Moreover, the high expression level of hMSL2 mRNA in Thymus and T-cells (Supplemental Figure 2) suggests a possible involvement in V(D)J recombination, a process that shares the machinery [38]. For these reasons we questioned whether Msl2 participates in NHEJ.

We first used in vivo end-joining reporter assays, whereby GFP cDNA encoded in a plasmid is blunt-digested with the restriction enzyme Xmn1, and transiently transfected into cells. The cell’s ability to re-ligate the broken DNA is measured by the level of GFP protein expressed, as judged by flow cytometry analysis. DT40 cells lacking Msl2 had an impaired ability (~25–50%) to re-ligate the digested DNA compared to wild-type cells (Figure 3A). A knockout cell line lacking Pfds, the chicken orthologue of DNA-PKcs [39], an essential component of the NHEJ repair pathway [40], also showed a decreased ability (71%) to re-ligate the DNA in this assay. As this protein is a crucial factor for NHEJ, we expected a larger defect in this assay for this cell line. We presume that this anomaly is due to limitations in the transfection efficiency in the DT40 system. We therefore wanted to test whether this defect was present in other systems. We first used siRNAs to deplete hMSL2 in U2OS cells. Depletion of the majority of hMSL2 was verified by western blotting analysis using a novel monoclonal antibody raised against a fragment of hMSL2 (Figure 3B, Supplemental Figures 1, 3). We then used the assay described above to test the ligation efficiency of these cells. Again we found that cells depleted of hMSL2 had an impaired ability (72%) to repair the digested DNA compared to control siRNA treated cells (Figure 3C).

We also utilized an established intrachromosomal NHEJ substrate-based system [41]; whereby two specific breaks are induced in an integrated reporter cassette by expression of the I-SceI restriction enzyme. Joining of the broken DNA through an NHEJ mechanism can be measured by expression of cell-surface markers using flow cytometry. Depletion of hMSL2 from cells in this assay resulted in a reduction (58%) in the frequency of end-joining compared to control siRNA treated cells (Figure 3D). Interestingly, a similar reduction (62%) was observed in cells that were depleted of hMOF (Figure 3D).

Taken together, these results implicate Msl2/hMSL2 in the NHEJ-mediated repair of DNA damage.

Msl2/hMSL2 is Stabilized in Response to DNA Damage

An obvious next question to address is how Msl2 behaves in response to DNA damage. In DT40 cells we could not monitor endogenous Msl2, as neither the monoclonal anti-hMSL2 antibody, nor three commercially available antibodies recognize
the chicken protein, therefore the Msl2-rescue cell line was used instead. We exposed these cells to 5 Gy of ionizing γ-irradiation (IR), then monitored the cells over a timecourse of 12 hours (Figure 4A). We used the phosphorylation of H2AX on serine 139 (γH2AX) to indicate the DNA damage response over this timecourse; levels begin to increase after 30 minutes, peak after

Figure 2. Histone modifications are perturbed by loss of Msl2. (A) Representative immunoblot analysis of nuclear extracts prepared from wild-type, Msl2<sup>−/−</sup>, and Msl2-rescue cell lines. Panels were probed with the antibodies indicated. (B) Quantification of (A). Mean expression levels of the various modifications in the cell lines were quantified and expressed relative to those in wild-type cells after normalization to H3 levels. Error bars represent standard deviation (n ≥ 3). (C) Q-PCR showing expression levels of the genes indicated. Expression levels in the cell lines are expressed relative to those in wild-type cells, following normalization to Gapdh. doi:10.1371/journal.pone.0068549.g002

Figure 3. Cells lacking Msl2/hMSL2 have defects in NHEJ repair. (A) End-joining efficiency in wild-type, Msl2<sup>−/−</sup> and Prkdc<sup>−/−</sup> DT40 cell lines as determined by GFP expression measured by flow cytometry analysis, following 16 hour transfection of Xml1-digested GFP plasmid. Transfection efficiency was normalized using uncut GFP plasmid. Repair efficiency was compared to wild-type cells. Compared to wild-type (n = 5): p value of Msl2<sup>−/−</sup> #1 = 0.0237 (n = 5); p value of Msl2<sup>−/−</sup> #2 = 0.0586 (n = 3); and p value of Prkdc<sup>−/−</sup> = 0.0912 (n = 5). (B) Representative immunoblot showing hMSL2 (upper two panels) and hMOF (lower two panels) depletion achieved using hMSL2 and hMOF siRNA respectively. (C) End-joining repair assay as in (A) in U2OS cells treated with control (cont)- or hMSL2-siRNAs. p value = 0.0014 (n = 4). (D) NHEJ repair efficiency in GC92 cells treated with control-, hMSL2- or hMOF-siRNAs as determined by repair of I-SceI digested intrachromosomal reporter. Quantified as percentage of CD4 positive cells by flow-cytometry analysis. Efficiency is compared to control cells. p value of hMSL2-siRNA treated cells = 0.0318 (n = 4) and p value of hMOF-siRNA treated cells = 0.1947 (n = 2). Error bars represent standard deviation. Two-tailed Student’s t-test was used to generate p values. doi:10.1371/journal.pone.0068549.g003
three hours, and return to pre-damage levels by nine hours. Anti-Flag antibody was used to detect the amount of HA2F-Msl2 protein. Interestingly, HA2F-Msl2 protein appeared to accumulate under these conditions; levels begin to increase after one hour and peak at three hours, before returning to pre-damage levels after nine hours. This increase is due to some post-transcriptional effect as there was no significant change in Msl2 mRNA levels over the timecourse (Figure 4B).

To investigate this stabilization further, we treated the DT40 Msl2-rescue cells with the proteasome inhibitors ALLN (N-Acetyl-L-leucyl-L-leucyl-L-norleucinal), and MG132. Both treatments caused an increase in abundance in the amount of HA2F-Msl2 protein as detected by anti-Flag antibody (Figure 4C). Intriguingly, alongside the accumulation of HA2F-Msl2, we noticed the appearance of a band slightly larger than the endogenous protein, suggestive of some post-translational modification.

We next examined endogenous hMSL2 in human U2OS cells. Cells were treated with 10 Gy IR and followed with a timecourse of 24 hours (Figure 4C). Again, we found hMSL2 protein accumulating; beginning 15 minutes after IR, and peaking after three hours. This accumulation followed slightly behind γH2AX detection in this case, which began after five minutes and peaks after 30 minutes. Alongside the accumulation of hMSL2, we again noticed the appearance of a band slightly larger than the endogenous protein. Also in keeping with the DT40 result (Figure 4C) is the accumulation of hMSL2 protein when cells are treated with ALLN (compare first and last lanes, Figure 4C).

These results suggest that soon after, or coincident with the formation of γH2AX foci, Msl2/hMSL2 accumulates in the cell. This accumulation is possibly due to the inhibition of proteolysis of Msl2/hMSL2 by the proteasome, limiting its turnover.

hMSL2 Mediates Modification of 53BP1 at Lysine 1690

hMSL2, in cooperation with hMSL1, ubiquitylates histone H2B on lysine 34 (H2BK34ub) [24]. This modification has links with transcriptional regulation, but not with the DNA damage response. In an attempt to identify other DDR-relevant substrates of hMSL2 we blasted the sequence surrounding H2BK34. A number of hits were returned; the first seven relevant (containing the lysine corresponding to K34) protein hits are listed in Table 1. Surprisingly, two of these were located in 53BP1 (Figure 5A). This information prompted us to investigate whether 53BP1 is a substrate of hMSL2.

We used a construct consisting of the minimal domain of 53BP1 required for foci formation (M-domain) [42] to generate V5-tagged M-domain constructs containing lysine to arginine point mutations of the residues that correspond to the predicted residues from the blast (K1568 and K1690), as well as point mutation of lysine 1273, known to be ubiquitylated by RAD18 [43]. A triple mutant of these residues was also generated (KtripleR). These constructs were then co-transfected with hMSL2 into U2OS cells for 24 hours and then analysed by immunoblotting analysis. WT, K1273R and K1568R constructs were modified in the presence of exogenous hMSL2 (Figure 5B). However, neither the K1690R construct nor the triple mutant showed the extra band (Figure 5B). We hypothesize that 53BP1 is monoubiquitylated at lysine 1690 by hMSL2.

We also wanted to test whether depletion of hMSL2 caused defects in 53BP1 recruitment to damage foci following IR treatment. However, under the conditions used, any differences observed between control and hMSL2-depleted cells were not significant.
DNA Damage-enhanced Modification of hMSL1 and hMOF via hMSL2

In addition to ubiquitylation of histone H2BK34 [24], it was recently reported that Drosophila MSL2 can ubiquitylate other components of the MSL complex including MOF, MSL1, and MSL3 [18]. We therefore questioned whether hMSL2 could ubiquitylate other members of the human MSL complex, and whether this could contribute to the DNA damage response. U2OS cells were transfected with HA2F-hMSL2, followed by treatment with IR. In the presence of HA2F-hMSL2 we detected an extra band (indicated by open arrowhead) above endogenous hMSL1 (Figure 5C) and hMOF (Figure 5D). This additional band is even more pronounced 15 minutes after treatment with 10 Gy IR. The shift in size is suggestive of mono-ubiquitylation. We therefore hypothesize that both hMSL1 and hMOF are ubiquitylated by hMSL2 in response to DNA damage.

Discussion

Despite its essential role in Drosophila dosage compensation, human MSL2 is relatively poorly characterized. Dosage compensation in mammals is mediated by a different mechanism and by different players [44], so it is likely that hMSL2 has some other function(s). Loss of function approaches using siRNA/shRNA mediated depletion of hMSL2 are difficult; levels of the protein vary between cell lines (Supplemental Figure S3), and depletion requires several rounds of siRNA treatment that does not always achieve acceptable reduced levels of protein [24] (and data not shown). To completely ablate Msl2 and better understand its function, we generated and characterised novel DT40 cell lines in which the Msl2 gene has been disrupted.

These Msl2 knockout cells appear normal, with only a slight growth defect (Figure 1D). This is surprising in light of our finding that several important histone modifications are disrupted in these cells. However, none of the modifications tested were completely lost, with levels reduced to between ~25–50% that in wild-type cells (Figure 2B). These residual levels may be sufficient for the cells to grow almost normally. This is also consistent with our proposal of Msl2’s role in the DNA damage response; without damage we would not expect to see a major growth phenotype.

The stabilization of Msl2/hMSL2 in response to damage strongly supports the proposed role of this protein in the DDR (Figure 4). As ALLN/MG132 treatment also leads to hMSL2/Msl2 accumulation, this stabilization appears to be dependent on the avoidance of some proteolytic action against hMSL2/Msl2 that may be keeping it at a basal level.

Using two standard assays to measure NHEJ efficiency [5], we have found that both chicken Msl2 and human MSL2 contribute to end-joining repair (Figure 3). The requirement of Msl2/hMSL2 for a fully functional NHEJ repair pathway is complex but
probably partly converges at the recruitment of 53BP1, a key protein in the NHEJ pathway [5,6,7]. 53BP1 is recruited to DNA damage sites by different means: acetylation of H4K16 [28]; dimethylation of H3K79 [8,11]; and di-methylation of H4K20 [9,11,37] have all been shown to (directly or indirectly) recruit 53BP1. We have observed reduced levels of all these modifications (Figure 2) in the Msl2 knockout cells. The reduction in H4K16ac may be due to lack of MSL complex formation/stability in the absence of Msl2/hMSL2 as previously described in Drosophila [16,45], or, as hMSL2 possibly ubiquitylates hMOF (Figure 5D), as was recently shown in Drosophila [46], it is conceivable that this hypothetical ubiquitylation promotes hMOF’s activity towards H4K16. Indeed, as hMOF and H4K16ac are known to have a role in the DNA damage response [28,29,30,31], it is possible that the defects in NHEJ observed here are largely due to aberrant hMOF activity in the absence of hMSL2 regulation.

The reduction in H3K79me2 is probably due to presumed loss of hMSL2 mediated H2BK34 ubiquitylation and the subsequent loss in stimulation of the DOT1L methyltransferase as previously reported [24]. Alongside the reduction in H3K79me2, this last study found reduced H3K4me3 in hMSL2-depleted cells [24]. In yeast, H3K4me3 was demonstrated to be involved in NHEJ [47]. We did not check H3K4me3 levels in our system.

In addition to histone modifications, 53BP1 also gets recruited to damage sites by interaction with the mediator protein MDC1 [20]. The hypothetical ubiquitylation of this residue is part of the nuclear localization sequence. It is noteworthy that Nucleoporin 153 (NUP153), a component of the nuclear pore [50] was recently shown to promote the nuclear import of 53BP1 important for the DDR [51]. Furthermore, NUP153 and hMSL2 are components of the hMSL/hMOF complex [20]. The hypothetical ubiquitylation of this residue could affect 53BP1 interaction with NUP153, and by extension the hMSL/hMOF complex. Alternatively, following initial recruitment to the aforementioned modifications, 53BP1 K1690 ubiquitylation could enhance its oligomerization, or could enhance its interaction with p33 or other proteins/modifications in such a way to promote the accumulation/function of this mediator protein at the site of damage.

We have shown that Msl2/hMSL2 plays a role in NHEJ, but it is possible that Msl2/hMSL2 plays a broader role in regulating the response to DNA damage. Higher levels of damage or unrepaired damage could lead to higher levels of hMSL2, resulting in ubiquitylation of p53, causing its nuclear export and the activation of the mitochondrial-dependent apoptotic pathway previously described [22,23]. Whereas, in response to low levels of damage we hypothesize that stabilized Msl2/hMSL2 could facilitate the histone/protein modifications described above, promoting the recruitment or accumulation of 53BP1 leading to NHEJ-mediated repair.

**Materials and Methods**

**Cell Culture**

Previously published DT40 cell lines [39], were provided by Ciaran Morrison (CCB, NUI Galway), and were cultured in RPMI media (Gibco) supplemented with 10% fetal calf serum (Lonza), 1% chicken serum (Sigma-Aldrich), and 1% penicillin/streptomycin (Sigma-Aldrich) at 39.5°C with 5% CO2. U2OS cells

![Figure 5. hMSL2 mediates modification of 53BP1, hMSL1 and hMOF.](image)

**Table 1. List of BlastP hits using H2B peptide as query.**

<table>
<thead>
<tr>
<th>KRGRKREYSI Blast Hit</th>
<th>Gene name</th>
<th>Max identity</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histone H2B type 3-8</td>
<td>H2B</td>
<td>100%</td>
<td>4e-05</td>
</tr>
<tr>
<td>Histone H2B type 1-8</td>
<td>H2B</td>
<td>90%</td>
<td>5e-04</td>
</tr>
<tr>
<td>Histone H2B type 1-M</td>
<td>H2B</td>
<td>80%</td>
<td>0.003</td>
</tr>
<tr>
<td>Sex comb on midleg-like protein 4</td>
<td>SCL4</td>
<td>70%</td>
<td>0.72</td>
</tr>
<tr>
<td>Dystonin</td>
<td>Bpag1</td>
<td>60%</td>
<td>0.73</td>
</tr>
<tr>
<td>Tumor protein p53 binding protein 1</td>
<td>53BP1</td>
<td>70%, 50%</td>
<td>1.9, 2.4</td>
</tr>
<tr>
<td>Ubiquitously transcribed tetratacipeptide repeat protein Y-linked</td>
<td>UTY</td>
<td>50%</td>
<td>2.6</td>
</tr>
</tbody>
</table>

The peptide KRGRKREYSI including K34 (underlined) in histone H2B was used as query in a search for homologous sequences in the human genome using the BLAST online tool. Listed are the first seven hits that include a lysine corresponding to K34. Also shown are their maximal identities and E-values.

doi:10.1371/journal.pone.0068549.t001
were commercially obtained from ATCC (American type culture collection), and ST4.5 cells [32] were provided by Rhodri Geredig (REMED1, NUI Galway), and both were cultured in DMEM (Dulbecco’s Modified Eagle’s Medium, Sigma) supplemented with 10% FBS (Fetal Bovine Serum, Sigma and Biosera) at 37°C with 5% CO₂. Cells were irradiated with gamma rays at the rate of 1294 Gy/hour using a Mainance Millennium 157Cs irradiator (Mainance Engineering Ltd). DT40 cells were treated with 50 or 100 μM of ALLN (N-Acetyl-L-leucyl-L-leucyl-L-norleucinal; Calbiochem) or 3 μM MG132 for 8 hours, while U2OS cells were treated with 100 μM of ALLN for 6 hours. Generation of Msl2<sup>−/−</sup> DT40 Cells

To disrupt the Msl2 gene, we generated Msl2-puromycin and Msl2-blasticidin disruption constructs by combining two genomic PCR products with the puromycin- or blasticidin-selection-marker cassette. The 2.4 kb targeting arm was amplified by PCR with primers 5′- GATCCTGTATCTTGGAGAGCC/TGTG-3′ and 5′- CAGATGTGAGTCAGTCTGCAAGAGAT-3′. The 4.1 kb targeting arm was amplified with primers 5′- actagTTGGATGAT-GAAATGTGCTGATAAATGT-3′ and 5′- aagctGAAATCTGCT-GAAGTGAACTGTGCTGCA-3′. Amplified PCR products were cloned into pGEM-T easy vector (Promega) and sequenced. To generate Msl2<sup>−/−</sup> cells, Msl2-puro and Msl2-bsr disruption constructs linearized with ApaLI were transfected sequentially by electroporation using the Gene Pulsar electroporation apparatus (Bio-Rad, Wicklow, Ireland). The genomic DNA of the transfectants was digested with BamHI, and gene-targeting events were confirmed by Southern blot using a probe external to the transgene. Of DT40 cells were treated with 100 μM of ALLN at 37°C with 5% CO₂. Cells were counted every 24h up to 96h.  

Cloning of Msl2 cDNA and Generate Stable DT40 Msl2<sup>−/−</sup> Rescue Cell Line

Chicken Msl2 cDNA was isolated by PCR amplification of the primary cDNAs using the 5′-ctcagatATGAAACCCGT-GAATGGCACA-3′ and 5′-gatccTCAAACGTCTATCTCA-GCTCATACGCT-3′ primers. The gene bank accession number of the chicken Msl2 gene is XM_426675. The PCR fragment was digested by XhoI and EcoRI and inserted into modified pCDNA3.1-HA-2xFlag vector. The result plasmid was used to generate stable DT40 Msl2<sup>−/−</sup> rescue cell line.  

Proliferation Analysis

For cell proliferation analysis, cultures were seeded in 24-well plates in triplicate at equal cell densities (5×10<sup>4</sup> cells/ml) and counted every 24h up to 96h. Flow Cytometry Analysis

Cells were harvested, fixed with 70% ethanol. For mitotic index determination, cells were treated with rabbit anti-H3ser10ph monoclonal antibody and subsequently with fluorescein isothiocyanate-conjugated anti-rabbit IgG antibody (Jackson Immunoresearch). The cells were resuspended in phosphate-buffered saline containing propidium iodide at 25 μg/ml and RNase-A at 250 μg/ml. The subsequent FACS analysis was performed with a FACS Canto apparatus and FACS Diva software (Becton Dickinson). For the analysis of the NHEJ assay, cells were washed in 1X PBS and DT40 cells were treated with 10 U of DNAseI for 15 minutes at room temperature. 5×10<sup>4</sup> of U2OS cells or 1×10<sup>5</sup> of DT40 cells were analysed for GFP expression using a FACS Canto (Becton Dickinson) and analysed using the BD FACSDiva Software (version 6.1.2, Becton Dickinson).

Quantitative Real-time PCR

Total RNA was obtained from DT40 cell lines using the ISOLATE RNA mini kit (Bioneer) and reverse transcribed using High capacity RNA-to-cDNA kit (Applied Biosystems), according to the manufacturers’ guidelines. cDNA was quantified following quantitative real-time PCR with primers against Msl2, Mof, Dot1l, Sax209, Gapdh, β-Actin using fast SYBR green master mix (Applied Biosystems) in a ABI 7500 fast (Applied Biosystems) according to manufacturers guidelines. Sequences of the real-time primers used are available upon request.

Antibodies and Immunoblotting

Whole-cell extracts were prepared with RIPA-buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM EDTA, and protease inhibitor cocktail). Nuclear and cytoplasmic fractions were obtained as previously described [53]. Proteins were transferred to nitrocellulose/PVDF membranes for analysis using the following primary antibodies: anti-H4K16ac (07–329, Upstate), anti-H4K20me2 (ab19164, Abcam), anti-H4K20me3 (07–463, Upstate), anti-H3K9me2 (#4653P, Cell Signaling Technology), anti-H3K79me2 (ab53594, Abcam), anti-H3ser10ph (06–570, Millipore), anti-Flag (F1804, Sigma), anti-beta-actin (ab9227, Abcam), anti-SCC1 [54], anti-alpha-tubulin (T 6074, Sigma) anti-gamma-H2AX (05–636, Millipore), anti-Flag (F1804, Sigma), anti-beta-actin (ab9227, Abcam), anti-SCC1 [54], anti-alpha-tubulin (T 6074, Sigma) anti-gamma-H2AX (05–636, Millipore), anti-Flag (F1804, Sigma), anti-beta-actin (ab9227, Abcam), anti-SCC1 [54], anti-alpha-tubulin (T 6074, Sigma) anti-gamma-H2AX (05–636, Millipore), anti-Flag (F1804, Sigma), anti-beta-actin (ab9227, Abcam), anti-SCC1 [54], anti-alpha-tubulin (T 6074, Sigma) anti-gamma-H2AX (05–636, Millipore), anti-Flag (F1804, Sigma), anti-beta-actin (ab9227, Abcam), anti-SCC1 [54]. Secondary antibodies: anti-mouse HRP (NA931, Abersham) and anti-rabbit-HRP (NA934, Abersham). For densitometry quantification, chemiluminescence was detected using a Fujifilm LAS3000 (Fujifilm) and quantified using ImageJ software.

siRNA Transfection

1×10<sup>6</sup> cells in a 6 well dish were treated with 10nM siRNA, 4 μl of oligofectamine (Invitrogen), in 1.5ml OptiMem (Invitrogen) for 18 hours. For hMsl2 depletion, cells received 3 rounds of treatment over 6–7 days. For hMOF, hMsl1 or hMsl3 depletion, cells were treated once for 2–3 days. The siRNAs used in this paper are as follows: control: siGENOME RISC-Free membrane (Dharmacon, sequence unavailable); hMSL2: ON-TARGETplus SMARTpool GUGUAUGCCAGGGAAACAA+ AAA-CAUCUAUAGCCGAAA+ CACCAUGCCUCGGCAAAU+ GUGUCUAAUUGAGGUGUA (Dharmacon), hMOF-G-CAAGAGAUCCGCGCAAA [dT][dT] (Sigma); hMof-G-CAAGAGAUCCGCGCAAA [dT][dT] (Sigma); hMof-G-CAAGAGAUCCGCGCAAA [dT][dT] (Sigma); hMof-G-CAAGAGAUCCGCGCAAA [dT][dT] (Sigma); hMof-G-CAAGAGAUCCGCGCAAA [dT][dT] (Sigma); hMof-G-CAAGAGAUCCGCGCAAA [dT][dT] (Sigma).

In vivo NHEJ Assay

U2OS cells were pre-treated siRNA, then 2×10<sup>5</sup> cells/well were seeded in a 6-well plate 24 hours before the assay. DT40 cells were grown until confluency (1×10<sup>6</sup> cells/ml) and 1×10<sup>5</sup> cells were used for the assay. U2OS cells were transfected with 1 μg of uncut pmaxGFP plasmid (Lonza) or with 1 μg of pmaxGFP plasmid linearized by restriction digest with XmnI enzyme within the GFP sequence using the Lipofectamine 2000 (Invitrogen). DT40 cells were electroporated with the same plasmids using the Amaxa GC92 cell line was performed as described previously [55].
Plasmid Transfection
U2OS cells were seeded in a 6-well plate (2 × 10^5 cells/well) 24 hours prior to plasmid transfection. Cells were transfected with 1 µg of indicated plasmids and cells were harvested 48 hours post-transfection directly in 2X Laemmli buffer. If subjected to IR treated, cells were harvested 15 minutes post-irradiation.

Supporting Information

Figure S1 Alignment of human hMSL2 and chicken Msl2. Msl2 was aligned with hMSL2 using bl2seq on NCBI. The RING domain is boxed in green, the CXC domain in red. A construct comprising residues 86 to 412 (sequence between RING and CXC domains) was used in the generation of the monoclonal anti-hMSL2 antibody. (TIF)

Figure S2 hMSL2 mRNA expression profile. Expression of hMSL2 mRNA in 79 human tissues according to the Affymetrix tool www.biogps.org [56,57]. Human U133A chip as analysed using the online bioinformatic tool www.biogps.org [36,57]. (TIF)

Figure S3 Generation and characterization of hMSL2 monoclonal antibody. (A) Coomassie stained gel showing induction and purification of 6His-hMSL2 fusion construct from Echerichia coli. 6His-hMSL2 comprises amino acids 86 to 412 of hMSL2 and has a predicted molecular weight (MW) of 45 kDa.

References

(B) Immunoblot analysis of whole cell extract from HeLa cells transfected with empty plasmid (mock) or a plasmid encoding HA-2F-hMSL2. Antibodies (hMSL2 hybridoma supernatant number) used are indicated below the blot. Endogenous hMSL2 has a predicted molecular weight of 75 kDa, and HA-2F-hMSL2, 80 kDa. 8A1, 4F12 and 8D2 correspond to different hybridoma supernatents tested. (C) Immunoblot analysis of U2OS cells transfected with siRNA against hMSL2 or with plasmids encoding HA-2F-hMSL2 or hMSL2-YFP (hMSL2 C-terminally-tagged with yellow fluorescent protein; MW 100 kDa). ST4.5 is a T-cell progenitor cell line [52]. (TIF)

Acknowledgments
We thank Brian McStay and Gaelle Legube for providing plasmids, Asifa Akhtar for anti-hMSL1 antibodies, Rhodri Ceredig for the T-cell line, Ciaran Morrison for DT40 cell lines, Noel Lowndes for antibodies, Tiago Dantas and Yifan Wang for reagents and advice concerning Msl2 targeting. We thank Ciaran Morrison and Noel Lowndes for critical reading of the manuscript.

Author Contributions
Conceived and designed the experiments: ZL, SM, YC, SR. Performed the experiments: ZL, SM, YC, SR. Analyzed the data: ZL, SM, YC, SR. Contributed reagents/materials/analysis tools: YC, SR. Wrote the paper: SR.

Msl2 in the DNA Damage Response

PLOS ONE | www.plosone.org 9 July 2013 | Volume 8 | Issue 7 | e68549


