Evidence of a MOF histone acetyltransferase-containing NSL complex in *C. elegans*

Matthew Hoe and Hannah R Nicholas*
School of Molecular Bioscience; University of Sydney; Sydney, Australia

Regulation of chromatin is a key process in the developmental control of gene expression. Many multi-subunit protein complexes have been found to regulate chromatin through the modification of histone residues. One such complex is the MOF histone acetyltransferase-containing NSL complex. While the composition of the human and *Drosophila* NSL complexes has been determined and the functions of these complexes investigated, the existence of an equivalent complex in nematodes such as *Caenorhabditis elegans* has not yet been explored. Here we summarise evidence, from our own work and that of others, that homologues of NSL complex components are found in *C. elegans*. We review data suggesting that nematode proteins SUMV-1 and SUMV-2 are homologous to NSL2 and NSL3, respectively, and that SUMV-1 and SUMV-2 may form a complex with MYS-2, the worm homolog of MOF. We propose that these interactions suggest the existence of a nematode NSL-like complex and discuss the roles of this putative NSL complex in worms as well as exploring the possibility of crosstalk between NSL and COMPASS complexes via components that are common to both. We present the groundwork from which a full characterization of a nematode NSL complex may begin.

**Introduction**

The regulation of chromatin is an important process in the developmental control of gene expression. Chromatin, composed of DNA and histone proteins, regulates the accessibility, and consequently the activity, of genes. Specific residues on the histone proteins can be modified, for instance through acetylation, methylation or phosphorylation. Modifications are made by chromatin modifying complexes, which are composed of multiple protein subunits including a histone modifying enzyme, such as a histone acetyltransferase (HAT), histone deacetylase (HDAC) or histone methyltransferase (HMT). One HAT of recent interest is KAT8/MOF (lysine (K) acetyltransferase/males absent on the first). KAT8/MOF is best known for its role in *Drosophila* dosage compensation as a component of the MSL (male-specific lethal) complex. More recently, however, it was found that KAT8/MOF is also a constituent of a separate complex called NSL (non-specific lethal). In *Drosophila*, the NSL complex was found to regulate constitutively expressed genes. ChIP-seq analysis revealed that it binds preferentially to promoters of housekeeping genes and that the NSL complex is required for efficient recruitment of RNA polymerase II to these promoters. The NSL complex is similarly found to regulate housekeeping genes in mouse embryonic stem cells (mESCs) and to also be required for the expression of key pluripotency factors in these cells.

Affinity purification and mass spectrometry analyses have revealed that the *Drosophila* and human NSL complexes are composed of 9 subunits: MOF, hMOF, dHCF/HCFC1, OGT/OGT1, WDS/WDR5, MCRS2/MCRS1, MBD-R2/PHF20, NSL1/KANSL1, NSL2/KANSL2 and NSL3/KANSL3 (Fig. 1). While the majority of these proteins have homologues in *C. elegans*, the existence of
a functional nematode NSL complex has yet to be determined. Here we describe each of the components of the NSL complex and summarise current knowledge of their homologues in *C. elegans* before discussing the possible functions of a putative nematode NSL complex.

### MOF

MOF is a member of the MYST family of HATs and is the enzymatic component of the NSL complex. Although in the context of the MSL complex MOF specifically acetylates histone 4 lysine 16 (H4K16ac),1,8 in the NSL complex MOF shows broader substrate specificity, additionally acetylating lysine 5 and 8 of H4.7 In *Drosophila*, the involvement of MOF in dosage compensation leads to a male-lethal phenotype in *mof* mutant flies,1 while murine MOF is required for both male and female embryos to develop beyond the expanded blastocyst stage.9 The *C. elegans* homologue of MOF, called MYS-2, also plays roles in development. For instance, reduction of *mys-2* function (in combination with reduction of function of another MYST HAT, *mys-1*) causes defects in cell fate maintenance.10

### HCFC1

Host cell factor C1 (HCFC1) is a transcriptional cofactor formed via the non-covalent association of multiple polypeptides, which are derived from a 2035 amino acid precursor protein through proteolytic cleavage (for a review see ref. 11). Initially identified through its interaction with transcription factor Oct-1 and the viral transactivating protein VP16 in the context of herpes simplex virus immediate early gene expression,12 HCFC1 has since been shown to interact with numerous cellular transcription factors. HCFC1 also associates with histone modification complexes in addition to NSL, including the Sin3 HDAC complex and the Set1/Asl2 HMT complex,13 suggesting that it may serve as a bridge between transcription factors and chromatin modulators.

The *C. elegans* homologue of HCFC1, called HCF-1, has been primarily studied in the context of longevity and stress responses. Interaction with HCF-1 is proposed to regulate the activity of the forkhead box O (FOXO) transcription factor DAF-16, a key longevity determinant, by sequestering a fraction of DAF-16 away from the promoters of its target genes.14 Additionally, at low temperatures, loss of *hcf-1* leads to smaller brood sizes and embryonic lethality due to defects in cell division, phenotypes which have been attributed to a reduction in histone 3 serine 10 phosphorylation (H3S10P).15

### OGT

O-linked *N*-acetylglucosamine (*O-GlcNAc*) transferase (OGT) catalyzes the *O-GlcNAcylation* of serine and threonine residues in a variety of nuclear and cytoplasmic proteins and it has been demonstrated that histone proteins are among the nuclear proteins that can be modified by OGT.16,17 The importance of OGT in mammalian development is evidenced by the finding that deletion of OGT compromises the viability of embryonic stem cells.18 While there are no obvious developmental defects associated with reduction of function of the nematode OGT, *C. elegans* OGT-1 has been implicated in the regulation of macronutrient storage since *ogt-1* mutants have increased glycogen levels and decreased lipid stores.19 Interestingly, in human and fly systems, OGT is responsible for the cleavage of the HCFC1 precursor into the subunits which associate to form the active protein.11,20 While the nematode HCF-1 protein, in contrast, does not appear to be subject to proteolytic processing,21 whether it is a substrate for O-GlcNAcylation has yet to be investigated.
WDR5

WDR5 is a WD40 repeat-containing protein. WD40 repeats are short motifs of approximately 40 amino acids that often end with a tryptophan and aspartic acid (WD) and are involved in protein-protein interactions (for a review see ref. 22). Three C. elegans homologues of WDR5 have been identified: WDR-5.1, WDR-5.2, and WDR-5.3.23,24 The best-characterized of these is WDR-5.1, which is required for germline stem cell maintenance and additionally plays a role in the transgenerational regulation of longevity.23,25

Like HCFC1, which interacts with many transcription factors and several histone modification complexes, WDR5 is not uniquely associated with the NSL complex. For instance, WDR5 is also a component of the COMPASS (Complex Proteins Associated with Set1) HMT complex. In this context, C. elegans WDR-5.1 has been found to be required for global di- and trimethylation of H3K4 in early worm embryos.23 Interestingly, coordinated activity of the NSL and COMPASS complexes has been proposed, with WDR5 acting as a bridge between the two.26 Recently, however, it has been reported that the interaction of WDR5 with NSL complex components occurs via the same regions with which WDR5 interfaces with COMPASS complex components, suggesting that the two interactions are mutually exclusive.27

MCRS1

Microspherule protein-1 (MCRS1) contains a forkhead-associated domain and a coiled-coil domain, both of which are protein-protein interaction motifs. Like HCFC1 and WDR5, MCRS1 has roles outside the NSL complex, for instance as a component of the INO80 chromatin remodeling complex.28 Human MCRS1 has been associated with a variety of cellular processes, including RNA polymerase I transcription and cell cycle progression.29,30 In Drosophila, MCRS2 is required for optimal recruitment of RNA polymerase II to the promoter regions of cyclin genes and consequently regulates cyclin gene expression.31 Although an MCRS1 homolog has been identified in C. elegans called MCRS-1, the functions of this nematode protein have not yet been examined.

PHF20

Plant Homeodomain (PHD) finger protein 20 (PHF20) is a multidomain protein incorporating two tudor domains, an AT hook motif and a PHD finger, for which it is named. The developmental importance of PHF20 is demonstrated by the perinatal mortality associated with loss of function of murine PHF20.32 Depletion of the Drosophila PHF20 homolog, MBD-R2, severely affects the mRNA levels of NSL target genes in a genome-wide manner.5 A C. elegans PHF20 homolog has not been identified.

The NSL Proteins

The NSL proteins NSL1, NSL2 and NSL3 appear to be unique to the NSL complex. In association with MOF, these proteins regulate housekeeping genes in Drosophila.23,25 NSL1, NSL2 and NSL3 are essential for Drosophila viability with lethality resulting from P-element insertions in the corresponding genes.6

The Drosophila NSL1 protein has been proposed to serve as a scaffold for the NSL complex, interacting directly with MOF, WDS, MCRS2 and MBD-R2.27 NSL1 is predicted to be intrinsically disordered, interacting with its partners through short linear motifs including the PEHE domain for interaction with MOF,33 and a WIN motif for interaction with WDS.27 We have not been able to identify an NSL1 homolog in the nematode proteome. However, since detection of homology for unstructured proteins such as NSL1 is challenging, the existence of a nematode equivalent of NSL1 cannot be excluded at this stage.

Although there is no clear nematode homologue of NSL1, in a recent publication we reported the identification of homologues of NSL2 and NSL3.34 We named these proteins SUMV-1 (C34E10.8) and SUMV-2 (F54D11.2) (suppressor of multivulva), respectively, because mutation of the genes encoding either of these proteins suppresses the multiple vulva phenotype in a synthetic multivulva (synMuv) mutant background. The multiple vulva phenotype of synMuv mutants arises through the simultaneous mutation of two synMuv genes in different classes (e.g. synMuv A and synMuv B) (for a review of synthetic multivulva and synthetic multivulva suppressor genes, see ref. 35).

Similarity between C. elegans SUMV-1 and Drosophila NSL2 is apparent primarily in the N-terminal portion of the nematode protein (Fig. 2).34 Within this region, the nematode protein contains a domain called pfam13891, a putative C3HC3H double zinc finger motif annotated as a potential DNA binding domain of chromatin remodeling proteins and helicases,36 although experimental evidence supporting this classification is lacking. It is possible that domains belonging to pfam13891 are, like other small zinc-coordinating domains, instead involved in protein-protein interactions, potentially with modified histones.37,38 While a single pfam13891 domain is found in SUMV-1, NSL2 contains two of these domains, as does Ino80d, a subunit of the INO80 chromatin remodeling complex.28 Recently a WDR5-interacting motif has been identified in NSL227 and we have found a similar motif in SUMV-1 (core sequence DEIDLL, Fig. 2), suggesting that this interaction may be conserved in nematodes.

Although similarly between C. elegans SUMV-2 and Drosophila NSL3 is only evident in the C-terminal portion of the nematode protein (Fig. 2),34 these proteins are classified as orthologous in OrthoList, as judged by orthology prediction programs InParanoid and Ensemble Compara.39 Within the region of similarity, both proteins are predicted to have an α/β-hydrolase fold.40

A putative nematode NSL-like complex

Prior to our work identifying similarity between SUMV-1 and SUMV-2 and NSL2 and NSL3, respectively, the two nematode genes sumv-1 and sumv-2 had been identified in a genome-wide RNA
interference (RNAi) screen for synMuv suppressor genes. Interestingly, this screen also identified C. elegans homologues of NSL complex components WDR5 (wdr-5.1) and HCFC1 (hcf-1) as synMuv suppressors. In our study, we additionally showed synMuv suppression with knockdown of mys-2, the worm homologue of MOF. That 5 genes which encode proteins that are homologous to components of the Drosophila and mammalian NSL complexes show this same reduction-of-function phenotype, suggests that the proteins that they encode are functionally related. We suggest therefore, that these proteins may act together within a nematode NSL-like HAT complex.

As an initial assessment of this hypothesis, we performed yeast two-hybrid assays and found that SUMV-1 and SUMV-2 interact in this system. We also examined whether these two proteins could interact with MYS-2 and, although we detected no interaction between SUMV-1 and MYS-2, we did find that SUMV-2 was a binding partner of MYS-2. Further assays will be required to determine whether there is a physical association between the other putative complex components including HCF-1, WDR-5.1, MCRS-1 and OGT-1, and to test whether reduction of mcrs-1 or ogt-1 function gives rise to a synMuv suppression phenotype, as would be expected if these proteins are functionally linked in an NSL-like complex.

In addition to identifying the genes encoding putative NSL complex components as synMuv suppressor genes, both our study and the earlier RNAi screen also found that RNAi against dpy-30 suppressed the synMuv phenotype. The DPY-30 protein is homologous to human protein hDPY-30, which is a member of the COMPASS complex. In light of the reported co-ordination of NSL and COMPASS complex activity in humans noted above, it is of interest that we also demonstrated a direct interaction between Dpy-30 and both SUMV-1 and SUMV-2 using yeast two-hybrid assays. The common phenotype of synMuv suppression, combined with evidence of physical association, suggests that DPY-30 might facilitate a functional interaction between the nematode COMPASS complex and putative NSL complex.

Possible roles of a putative NSL complex in C. elegans

Given that the NSL proteins appear to be uniquely involved in the NSL complex,
and since SUMV-1 and SUMV-2 are the only identified nematode homologues of the NSL proteins, the functions of these two proteins may shed light on the physiological roles of the putative \textit{C. elegans} NSL complex. Beyond vulval specification, a further developmental role for this complex is suggested by two observations pertaining to SUMV-2. During normal \textit{C. elegans} development, germ-line differentiation is suppressed in somatic cells. However, if the function of a protein called MEP-1 is compromised, the germ-line-soma distinction is lost, with the result that germ-line genes such as \textit{P-granule components} are ectopically expressed in the soma. \textsuperscript{43} This loss of somatic suppression of germ-line patterns of gene expression is accompanied by larval lethality. \textsuperscript{44} suggesting that SUMV-2 antagonises the activity of MEP-1 in maintaining the germ-line-soma distinction. \textit{mep-1} is a synMuv B gene and, interestingly, loss of function of several other synMuv B genes similarly results in ectopic somatic expression of germ-line genes but is not accompanied by larval arrest under standard conditions. \textsuperscript{45} However, when the function of these other synMuv B genes is compromised at high temperature, the ectopic somatic expression of germ-line genes is enhanced and this is accompanied by larval lethality, with the latter referred to as High Temperature Arrest (HTA). \textsuperscript{35} Consistent with a role for SUMV-2 in antagonising the activity of the synMuv B genes in the maintenance of the germ-line-soma distinction, \textit{sumv-2} RNAi suppression of \textit{mep-1} mutants grown under standard culture conditions. RNAi of \textit{sumv-2} partially rescues the lethality of \textit{mep-1} mutants, \textsuperscript{41} suggesting that SUMV-2 antagonises the activity of MEP-1 in maintaining the germ-line-soma distinction. \textit{mep-1} is a synMuv B gene and, interestingly, loss of function of several other synMuv B genes similarly results in ectopic somatic expression of germ-line genes but is not accompanied by larval arrest under standard conditions. \textsuperscript{44} However, when the function of these other synMuv B genes is compromised at high temperature, the ectopic somatic expression of germ-line genes is enhanced and this is accompanied by larval lethality, with the latter referred to as High Temperature Arrest (HTA). \textsuperscript{35} Consistent with a role for SUMV-2 in antagonising the activity of the synMuv B genes in the maintenance of the germ-line-soma distinction, \textit{sumv-2} RNAi suppresses the HTA phenotype of several synMuv B mutants. \textsuperscript{45}

There is also evidence implicating SUMV-1 and SUMV-2 in the regulation of transgene expression in \textit{C. elegans}. The expression of multiple-copy transgenes is normally silenced in the nematode germ-line and in somatic cells lower levels of expression are detected per gene copy from such transgenes than from the corresponding endogenous gene. \textsuperscript{46,47} A \textit{let-858::GFP} transgene is usually expressed in somatic tissues but silenced in the germ-line, however, knockdown of \textit{sumv-2} resulted in germ-line expression of this transgene, \textsuperscript{41} suggesting that SUMV-2 is required for germ-line transgene silencing. In addition, we observed that reduction of \textit{sumv-1} function increased hypodermal expression from a multiple-copy \textit{cbrp-1::gfp} transgene, which may indicate a role for SUMV-1 in the somatic regulation of transgene expression. \textsuperscript{34}

SUMV-1 and SUMV-2 may also regulate expression of the Notch ligand \textit{LAG}-2. While a number of synMuv genes repress expression of a \textit{lag-2::gfp} reporter, such that ectopic expression of this reporter is observed in the intestinal cells (among others) of synMuv B mutant animals. \textsuperscript{48,49} knockdown of either \textit{sumv-1} and \textit{sumv-2} in a synMuv B mutant background reduced the proportion of animals with intestinal \textit{lag-2::GFP} expression. \textsuperscript{41} Thus SUMV-1 and SUMV-2 are required for the ectopic expression of \textit{lag-2} and may serve as transcriptional activators of this gene.

Finally, SUMV-1 and SUMV-2 play a role in RNAi. This role was determined in the context of depletion of a gene called \textit{mom-2} by RNAi, which usually gives rise to embryonic lethality. When \textit{mom-2} RNAi was performed concurrently with either \textit{sumv-1} or \textit{sumv-2} RNAi, a high proportion of surviving progeny was observed, \textsuperscript{41} suggesting that \textit{sumv-1} and \textit{sumv-2} are required for robust RNAi. Interestingly, although a role for the MYS-2 HAT in RNAi has not been examined using this assay, the importance of this protein in transgenerational inheritance of RNAi has been demonstrated. The effect of certain RNAi treatments in \textit{C. elegans} has been shown to be heritable for as many as 80 generations and this inheritance is dependent on MYS-2, among other factors. \textsuperscript{50}

**Conclusion**

The characterization of a putative NSL-like complex in \textit{C. elegans} has begun with our presentation of the first evidence of physical interactions between worm homologues of NSL2, NSL3 and MOF and with our work and that of others showing that putative complex components SUMV-1, SUMV-2, MYS-2, HCF-1 and WDR-5.1 play a common role in vulval development. However, further investigations are required to assess whether these proteins, together with OGT-1 and MCRS-1, genuinely function together in a HAT complex analogous to the \textit{Drosophila} and mammalian NSL complexes.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Funding**

This work was supported by a grant from the Australian Research Council awarded to H.R.N and by an Australian Postgraduate Award awarded to M.H.

**References**

1. Hilfiker A, Hilfiker-Kleiner D, Pannuti A, Lucchesi JC. mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in \textit{Drosophila}. EMBO J 1997; 16:2854-60; PMID:9155031; http://dx.doi.org/10.1093/emboj/16.8.2854

2. Lam KC, Muhlpfordt F, Vaquerizas JM, Raja SJ, Holt H, Luscombe NM, Manke T, Akhtar A. The NSL complex regulates housekeeping genes in \textit{Drosophila}. PLoS Genet 2012; 8:e1002736; PMID:22723752; http://dx.doi.org/10.1371/journal.pgen.1002736

3. Ravens S, Fournier M, Ye T, Sterle M, Demble D, Chavant V, Tora L. Mof-associated complexes have overlapping and unique roles in regulating pluripotency in embryonic stem cells and during differentiation. eLife 2014; 3; PMID:24898753; http://dx.doi.org/10.7554/eLife.02104

4. Chelmiicki T, Dundar F, Turley MJ, Khanam T, Akbas T, Ramirez F, Gendrel AV, Wright PR, Vedom P, Bucko-Fen R, et al. Mof-associated complexes ensure stem cell identity and Xist repression. eLife 2014; 3; e02024; PMID:24842875; http://dx.doi.org/10.7554/eLife.02024

5. Raja SJ, Carapin I, Conrad T, Vaquerizas JM, Gebhardt P, Holt H, Kadile J, Fraterman S, Luscombe NM, Akhtar A. The non-specific lethal complex is a transcriptional regulator in \textit{Drosophila}. Mol Cell 2010; 38:827-41; PMID:20620954; http://dx.doi.org/10.1016/j.molcel.2010.05.021

6. Mendjan S, Taipale M, Kind J, Holt H, Gebhardt P, Schelder M, Vermeulen M, Buscaino A, Duncan K, Mueller J, et al. Nuclear pore components are involved in the transcriptional regulation of dosage compensation in \textit{Drosophila}. Mol Cell 2010; 38:827-41; PMID:20620954; http://dx.doi.org/10.1016/j.molcel.2010.05.021

7. Cai Y, Jin J, Swanson SK, Cole MD, Choi SH, Florens A, Backofen R, et al. MOF-associated complexes ensure stem cell identity and Xist repression. eLife 2014; 3; e02104; PMID:24898753; http://dx.doi.org/10.7554/eLife.02104

8. Taipale M, Rea S, Richter K, Vilar A, Lichter P, Imhof A, Akhtar A. \textit{lMOF} histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. Mol Cell Biol 2005; 25:6798-810;
