Ectopic transcription due to inappropriately inherited histone methylation may interfere with the ongoing function of terminally differentiated cells

Juan D. Rodriguez[1], Monica Reeves[1] ,Hsiao-Lin Wang[2], Jaely Chavez[3], Rhea Rastogi[1], Sindy Chavez[1], Madhav Chadha[1], Emily J. Hill[2], Victor G. Corces[2], Karen Schmeichel[3], David J. Katz[1*]

[1]Department of Cell Biology, Emory University School of Medicine, Atlanta GA 30322, USA.
[2]Department of Human Genetics, Emory University School of Medicine, Atlanta GA 30322, USA.
[3]Department of Biology, Oglethorpe University, Brookhaven GA 30319, USA.

ABSTRACT

Many human neurodevelopmental disorders are caused by *de novo* mutations in histone modifying enzymes. These patients have craniofacial defects, developmental delay, intellectual disability and behavioral abnormalities, but it remains unclear how the mutations lead to such developmental defects. Here we take advantage of the invariant *C. elegans* lineage along with a unique double mutant in the H3K4me1/2 demethylase SPR-5/LSD1/KDM1A and the H3K9 methyltransferase MET-2/SETDB1 to address this question. We demonstrate that *spr-5; met-2* double mutant worms have a severe chemotaxis defect that is dependent upon the ectopic expression of germline genes in somatic tissues. In addition, by performing single-cell RNAseq, we find that germline genes begin to be ectopically expression widely in *spr-5; met-2* embryos. However, surprisingly we found that *spr-5; met-2* mutants have no somatic lineage defects prior to the 200-cell stage of embryogenesis. This suggests that the altered chemotaxis behavior may be due to ongoing defect in terminally differentiated cells rather than a defect in development. To test this directly, we used RNAi to shut off the ectopic expression of germline genes in L2 *spr-5; met-2* larvae, which have a fully formed nervous system. Remarkably, we find that shutting off the ectopic germline expression rescues normal chemotaxis behavior in the same adult worms that previously had a chemotaxis defect at the L2 stage. This suggests that ongoing ectopic transcription can block normal behavior in a fully intact nervous system. These data raise the possibility that intellectual disability and altered behavior in neurodevelopmental syndromes, caused by mutations in histone modifying enzymes, could be due to ongoing ectopic transcription and may be reversible.

INTRODUCTION

Many human neurodevelopmental disorders are caused by de novo mutations in chromatin regulators. For example, Kabuki syndrome is caused by mutations in the H3K4 methyltransferase KMT2D and the H3K27 demethylase KDM6A[1]. Kabuki syndrome patients have developmental delay, craniofacial defects and intellectual disability[1]. Additionally, three patients have been identified with mutations in the H3K4me1/2 demethylase LSD1/KDM1A (referred to as SPR-5 in *C. elegans*). These LSD1 patients have been referred to as Kabuki-like because their characteristics are nearly identical to Kabuki Syndrome patients[2]. However, it remains unclear how defects in the regulation of histone modifications give rise to neurodevelopmental disorders, such as Kabuki Syndrome.

H3K4 methylation is acquired co-transcriptionally via the COMPASS complex interacting with RNA polymerase and is associated with actively transcribed genes[3,4]. H3K4 methylation may function as an epigenetic memory, maintaining transcription over time or through mitotic cell divisions[5,6]. As a result, H3K4 methylation that is acquired during the production of gametes may have to be erased to prevent this epigenetic memory from being propagated across generations. In *C. elegans*, SPR-5 is required maternally to erase H3K4me2 and prevent it from being inherited transgenerationally[7]. Without SPR-5, worms become increasingly sterile across generations (termed germline mortality) due to the transgenerational accumulation of H3K4me2 and the increasing expression of germline genes. Similarly, when LSD1 is mutated maternally in mice, it results in embryonic lethality at the 2-cells stage[8,9]. This demonstrates that SPR-5/LSD1 has a conserved role in maternal epigenetic reprogramming.

H3K9 methylation is associated with repressed transcription[10]. In *C. elegans*, loss of the H3K9 methyltransferase MET-2 (referred to as SETDB1 in mammals) also results in a germline mortality phenotype and *spr-5*; *met-2* double mutants have an exacerbated maternal effect sterility phenotype[11-13]. This suggests that SPR-5 and MET-2 cooperate in maternal epigenetic reprogramming. The function of MET-2 in maternal reprogramming is conserved, as maternal loss of SETDB1 in mice also results in early embryonic lethality[14].

In *C. elegans*, another transcription-associated histone modification, H3K36me3, is acquired via the H3K36 methyltransferase MET-1 during production of the gametes[15]. In the embryo, this H3K36 methylation is maintained by a transcription-independent H3K36 methyltransferase MES-4 at 176 critical germline genes[15,16]. For the remainder of the manuscript, we will refer to these genes as MES-4 targeted germline genes. At around the 60-cell stage, the germline blastomere P4 divides to give rise to the primordial germ cells, termed Z2 and Z3. Once the embryo hatches, the MES-4 targeted germline genes are expressed as Z2 and Z3 begin to proliferate and loss of MES-4 results in a maternal effect sterility phenotype[15,16]. Thus, it has been suggested that MES-4-dependent H3K36me3 may function as a type of bookmark to help re-specify the germline in the subsequent generation. The transcription factor LSL-1 also plays a critical role in germline function. LSL-1 is first transcribed in P4, continues to be expressed during germline development in all four stages of larval development (L1 to L4), and remains on into the adult. Without, LSL-1 the germline has many defects including defects in meiosis, germline apoptosis and the production of almost no functional gametes[17].

spr-5; met-2 double mutants have a severe developmental delay at the L2 larval stage that is associated with the ectopic expression of MES-4 targeted germline genes in somatic tissues. In addition, *spr-5; met-2* mutants have ectopic H3K36me3 at MES-4 targeted genes in somatic tissues. When mes-4 is knocked down by RNA interference in *spr-5; met-2* mutants, this ectopic expression is eliminated and the L2 larval delay is rescued, suggesting that the severe developmental delay in *spr-5; met-2* mutants is caused by the ectopic expression of MES-4 targeted genes[18]. The ectopic maintenance of H3K36me3 in *spr-5; met-2* mutants is likely propagated by H3K4 methylation, because the developmental delay of *spr-5; met-2* mutants is also dependent on the H3K4 methyltransferase SET-2[18]. Similar to *spr-5; met-2* mutants, loss of a NuRD complex component MEP-1 or DREAM complex component LIN-35 also results in the ectopic expression of germline genes and this ectopic expression is also dependent on MES-4[19,20]. This suggests that these complexes may be functioning to reinforce SPR-5/MET-2 maternal reprogramming in somatic tissues. In the germline, the function of NuRD is antagonized by the germline transcription factor LSL-1[17].

In *C. elegans*, the embryonic lineage is completely invariant. This means that every wild-type embryo undergoes the same pattern of cell division and cell migration[21]. The invariant cell lineage of *C. elegans* enabled the discovery of programmed cell death (apoptosis)[22]. In order to facilitate study of the invariant lineage, the Waterston lab developed automated lineage tracing. The system utilizes a ubiquitously expressed histone-mCherry fusion protein to track cells via their nuclei with 3D time-lapse confocal imaging[23]. In *spr-5; met-2* double mutants, there is a massive accumulation of H3K4me2 and the ectopic expression of MES-4 targeted germline genes[11,18]. The

automated lineage tracing system, along with the invariant *C. elegans* lineage, provide the unique opportunity to understand how the inappropriate inheritance of chromatin and the ectopic expression of genes affects cells at the single cell level and gives rise to phenotypes in *spr-5; met-2* mutants.

To determine what embryonic cell types the MES-4 targeted germline genes are ectopically expressed in and how this may alter the embryonic lineage, we performed single-cell RNAseq and automated lineage tracing on *spr-5; met-2* mutants. We find that MES-4 targeted germline genes begin to be ectopically expressed broadly in many embryonic lineages, except in the germline itself where MES-4 targeted germline genes may not be fully activated. Surprisingly, the ectopic expression of MES-4 targeted germline genes in somatic lineages does not result in any somatic lineage defects through the 200-cell stage of embryogenesis. This raises the possibility that the ectopic expression of MES-4 germline genes does not alter development, but instead causes an ongoing defect in differentiated cells. While performing these experiments, we noticed that *spr-5; met-2* mutants fail to move toward OP50 bacteria. Using chemotaxis assays, we found that *spr-5; met-2* mutants have a chemotaxis defect that begins at the L2 stage and is dependent upon MES-4, suggesting that it is dependent on the ectopic expression of MES-4 targeted germline genes. This observation provided the opportunity to test whether the ectopic expression of MES-4 targeted germline genes may be actively interfering with the function of differentiated cells by determining whether blocking the ectopic expression of these genes restores normal chemotaxis. Remarkably, we find that shutting off the ectopic expression of the MES-4 after the L2 stage rescues normal chemotaxis behavior. From these data, we conclude that the

ectopic expression of MES-4 targeted germline genes can alter normal behavior in a fully intact nervous system.

RESULTS

Germline genes are ectopically expressed in the somatic lineages of *spr-5; met-2* mutant embryos

Previously we found in *spr-5; met-2* mutants that MES-4 targeted germline genes are ectopically expressed in the soma at the L1 stage[18]. To determine how these genes are misexpressed in the embryo at the single-cell level, we performed single-cell RNA sequencing at the 100-160 cell stage. At this stage gastrulation is completed and zygotic transcription has fully initiated[24,25]. We obtained sequences from 686 *spr-5; met-2* mutant cells and 219 Wild Type (N2) cells and identified significantly changed genes using a 0.25-fold change cutoff. *Seurat* clustering analysis of the cells from both Wild Type and *spr-5; met-2* mutants identified 8 clusters (Figure 1A). The individual cells from both Wild Type and *spr-5; met-2* mutants fall within the same clusters, suggesting that gene expression is not dramatically changed overall in *spr-5; met-2* mutants compared to Wild Type (Figure 1B). Consistent with SPR-5 and MET-2 acting as transcriptional repressors, there are 3,660 genes significantly upregulated and only 1,679 genes significantly downregulated in our single-cell dataset.

To determine whether MES-4 targeted germline genes are ectopically expressed in each cluster, we examined the gene expression (normalized count matrices) of all MES-4 targeted germline genes in each cluster. In all 8 clusters, the average gene expression across all MES-4 targeted germline genes is higher in *spr-5; met-2* mutants compared to Wild Type. We also determined the percentage of cells within each cluster

that express each of the MES-4 targeted germline genes (Figure 1C-J). In 5 of the 7 somatic clusters (0,1,4,5,6) there is a large increase the percentage of cells within the cluster that express MES-4 targeted germline genes (Figure 1C-G). Consistent with this, 25 of the 149 examined MES-4 targeted germline genes are significantly ectopically expressed in one of the 7 somatic clusters (Figure 1K). Together these results suggests that MES-4 targeted germline genes begin to be widely expressed in many somatic lineages in the embryo.

Despite the ectopic expression of MES-4 targeted germline genes in *spr-5; met-2* mutants overall, in the germline cluster (7) as well as the hypodermis and body wall muscle cluster (2) the percentage of cells within the cluster that express MES-4 targeted germline genes decreases in *spr-5; met-2* compared to Wild Type (Figure 1H,I). Consistent with this decrease in the percentage of cells expressing MES-4 targeted genes in the germline, there are only 20 MES-4 targeted germline genes significantly decreased in any cluster and 16 of them are in cells within the germline cluster (7). However, the decrease in percentage of cells expressing MES-4 targeted genes in clusters 7 and 2 occurs despite there still being an increase in the average gene expression of the MES-4 targeted germline genes, as there is in all clusters. Overall, the decrease in the expression of some MES-4 targeted germline genes in the germline cluster (7) is consistent with the possibility that many MES-4 germline genes fail to activate in the germline of *spr-5; met-2* mutants. It is not entirely clear why MES-4 targeted germline genes are also expressed within fewer cells within cluster 2, but this effect appears to be driven by a higher percentage of these cells expressing MES-4 germline genes in Wild Type. Thus, it is possible that the higher level of expression of

MES-4 germline genes that occurs normally in cluster 2 has to do with many of the cells within cluster 2 being immediately adjacent to the germline cluster (7). In the excretory, glia and seam somatic cell cluster (3), the percentage of cells within the cluster that express MES-4 targeted germline genes is unchanged between N2 and *spr-5; met-2* mutants (Figure 1J). This raises the possibility that cells of this lineage may be resistant to the ectopic expression of MES-4 targeted germline genes.

To confirm that MES-4 targeted genes are ectopically expressed in *spr-5; met-2* mutants, we performed quantitative RT-PCR on 5 of the MES-4 targeted genes that were not significantly misexpressed in our single-cell dataset but were significantly misexpressed at the L1 stage in *spr-5; met-2* mutants[18]. We found that 4 out of the 5 genes tested were significantly upregulated with an average fold change of 3.5 (Supplemental Figure 1). Consistent with these results, we previously found by single molecule RNA fluorescence *in situ* hybridization that one of the genes, *htp-1,* is widely ectopically expressed in the somatic cells of the embryo[18]. Finally, we also compared the ectopic expression of MES-4 targeted germline genes in our single-cell RNAseq dataset to our previously published *spr-5; met-2* RNAseq performed at the L1 larval stage. In contrast to our single cell dataset, where only 25 MES-4 targeted germline genes are significantly upregulated each only in a single cluster, in our previous L1 RNAseq dataset there were 34 MES-4 targeted genes that were significantly upregulated across the entire L1 larvae. This suggests that MES-4 targeted germline genes are increasingly ectopically expressed at later stages. Overall, our scRNAseq data suggest that inappropriate chromatin inherited from the previous generation in *spr-*

5; met-2 mutants is permissive for facilitating ectopic expression beginning widely during embryonic stages, but this ectopic expression is less than at later stages.

spr-5; met-2 mutants have a delay in the specification of the germline, but no defects in somatic embryonic development

To determine if cell lineage specification is altered in *spr-5; met-2* mutants, we took advantage of the invariant embryonic lineage in *C. elegans*. By performing automated lineage tracing[26], we tracked the number of cell divisions, timing of cell divisions and cell migration of all embryonic cells from the 2-cell stage through the 200-cell stage in *spr-5; met-2* mutants. Surprisingly, we found no defects in any somatic lineages deriving from either the AB or P1 blastomeres (Figure 2A-D). These data suggest that embryonic development is normal in *spr-5; met-2* mutants. In contrast, we identified a significant defect within the germline lineages (Figure 2B,C). Normally, the germline blastomere divides to give rise to the two primordial germ cells, Z2 and Z3, around the 60-cell stage, ~80 minutes after the 2-cell embryo first divides. In Wild Type, it takes an average of 45 minutes from the time that P_4 first appears to when P_4 divides to yield Z2 and Z3. In contrast, *spr-5; met-2* mutants require an average of 56 minutes (Figure 2C). The delay in the cell division of the germline blastomere P_4 occurs despite their being no delay in the cell D, directly adjacent to P_4 (Figure 2D). The P_4 delay correlates with the failure to fully express MES-4 targeted germline genes in the germline cluster (7)(Figure 1I), raising the possibility that the delayed division of P4 could be due to the failure to fully activate germline transcription.

spr-5; met-2 mutants have a severe defect in chemotaxis towards OP50 bacteria

While maintaining *spr-5; met-2* mutants and performing these experiments, we noticed that *spr-5; met-2* mutants fail to move towards the OP50 *E. coli* food source. To quantify this defect, we performed a chemotaxis assay and calculated the chemotaxis index (CI) (Figure 3). Compared to Wild Type worms which have a chemotaxis index of 0.9, *spr-5* and *met-2* single mutants have a small but significant defect in chemotaxis, with chemotaxis indices of 0.7. In contrast, *spr-5; met-2* mutants have a chemotaxis index of 0.2, indicative of a more severe defect in chemotaxis (Figure 3A). This chemotaxis defect is observed in both L2 larvae and adult animals (Figure 4). Importantly, in the chemotaxis assay, any animals that fail to move from the origin are not included in the analysis, so this defect is not due to a failure in mobility. Taken together, these results demonstrate that *spr-5; met-2* chemotaxis mutants fail to go towards the OP50 *E. coli* food source, despite not having any developmental defects in the embryonic lineage.

The _spr-5; met-2_ chemotaxis defect is dependent upon the ectopic expression of MES-4 targeted germline genes

Previously we demonstrated that a severe developmental delay also observed in *spr-5; met-2* mutants can be rescued by knockdown of the H3K36 methyltransferase MES-4[18]. This suggested that the developmental delay is dependent upon the ectopic expression of MES-4 targeted germline genes. To determine whether the chemotaxis defect of *spr-5; met-2* mutants is also dependent upon the ectopic expression of MES-4 germline genes, we measured chemotaxis in *spr-5; met-2* mutants in which *mes-4* is knocked down by RNA interference (RNAi) both maternally and zygotically. RNAi of *mes-4* in

spr-5; met-2 mutants significantly rescued the chemotaxis defect, suggesting that the chemotaxis defects is also dependent upon the ectopic transcription of MES-4 targeted germline genes (Figure 3B).

Zygotically knockdown of the ectopic transcription of MES-4 targeted germline genes in L2 *spr-5; met-2* mutants, restores normal chemotaxis

The lack of embryonic developmental defects in *spr-5; met-2* mutants raised the possibility that the severe chemotaxis defect in these mutants could be due to the ongoing ectopic expression of MES-4 targeted germline genes. If this is the case, it might be possible to shut off the ectopic expression of the MES-4 targeted germline genes and restore normal chemotaxis. To test this possibility, we took L2 *spr-5; met-2* larvae which already have a severe chemotaxis defect and attempted to shut off the ectopic expression of MES-4 targeted germline genes by performing *mes-4* RNAi. Importantly, because *spr-5; met-2* double mutants have a severe developmental delay, it takes ~5 days for them to develop from L2 larvae to adults. This gives the RNAi plenty of time to knockdown the ectopic transcription of the MES-4 targeted germline genes. Additionally, in these experiments the RNAi is performed with HT115 bacteria rather than OP50. This results in an overall reduction in the chemotaxis index. For example, under these conditions the chemotaxis index of Wild Type animals is 0.55 rather than 0.9 and the chemotaxis index of *spr-5; met-2* double mutants is 0.0 rather than 0.2. Remarkably, we find that shutting off the ectopic germline expression of the MES-4 targeted germline genes after the L2 stage results in a significant rescue of the chemotaxis index from 0.0 to 0.28 (Figure 4).

MES-4 is required to propagate the ectopic transcription of MES-4 targeted germline genes to somatic tissues but may not be completely required to continually maintain the ectopic transcription of MES-4 targeted germline genes in somatic cells. Therefore, we also tried to shut off the ectopic expression of MES-4 targeted germline genes by performing RNAi against *lsl-1.* LSL-1 is a germline transcription factor that is actively required for the transcription of many germline expressed genes[17]. We found that knockdown of LSL-1 after the L2 stage in *spr-5; met-2* mutants results in an even stronger rescue of the chemotaxis index in *spr-5; met-2* mutants from 0.0 to 0.36, such that the chemotaxis index is similar to Wild Type (0.55) under these experimental conditions (Figure 4). Taken together these results suggest that the ongoing ectopic transcription of MES-4 targeted germline genes can block normal chemotaxis behavior in a fully intact nervous system.

DISCUSSION

spr-5; met-2 double mutants inappropriately inherit chromatin from the previous generation and ectopically express germline genes in somatic tissues. This provided the unique opportunity to determine how the ectopic expression of genes when histone methylation is inappropriately inherited leads to phenotypes. Previously, we found that *spr-5; met-2* mutants have a severe L2 developmental delay that is associated with the ectopic expression of MES-4 targeted germline genes. Knock down of *mes-4* by RNA interference eliminates the ectopic expression of MES-4 targeted germline genes and rescues the L2 developmental delay, suggesting that the developmental delay is dependent upon the ectopic expression of MES-4 targeted germline genes[18]. Here we found that *spr-5; met-2* mutants also have a severe defect in chemotaxis towards OP50 bacteria at both the L2 larval stage and in adults. There is also a significant chemotaxis defect in both *spr-5* and *met-2* single mutants, but the chemotaxis defect is much more severe in *spr-5; met-2* double mutants, confirming the synergistic interaction between *spr-5* and *met-2*. In addition, we find that the severe chemotaxis defect in *spr-5; met-2* mutants is dependent upon MES-4, suggesting that this phenotype is also dependent upon the ectopic expression of MES-4 targeted germline genes.

Single-cell RNAseq analysis of *spr-5; met-2* mutants at the ~200 cell stage demonstrated the MES-4 targeted germline genes begin to be ectopically expressed in the embryo. However, the ectopic expression was less severe than what we previously observed at the L1 stage[18]. One potential explanation for the increased ectopic expression at the L1 stage is that the permissive chromatin must be acted upon by transcription factors that are not expressed until the individual lineages are further

differentiated. This could potentially explain why some phenotypes are not observed until later in development.

The ectopic expression of MES-4 targeted germline genes is observed equally in most somatic clusters. This is consistent with a model in which the ectopic chromatin is inherited fairly uniformly. However, we do not observe the ectopic expression of the MES-4 targeted germline genes in the excretory, glia and seam somatic cell cluster (3), so some lineages may not inherit the ectopic chromatin or may be resistant to the ectopic expression of these germline genes. Despite the ectopic expression of MES-4 germline genes, unsupervised clustering analysis of our *spr-5; met-2* single-cell RNAseq data demonstrated that these mutants have the same 8 clusters as Wild Type. This suggests that the ectopic transcription does not broadly interfere with cell specification in the embryo. Consistent with this possibility, our automated lineage tracing of *spr-5; met-2* mutants demonstrated that there are no obvious somatic defects in the embryonic lineage through the ~200 cell stage.

The lack of embryonic lineage defects occurs despite there being a severe developmental delay and impaired chemotaxis beginning at the L2 stage. These data are consistent with two possible models. One possibility is that embryonic development is resistant to the MES-4 targeted germline genes that are significantly ectopically expressed at the ~200 cell stage. In this scenario, the strong maternal programming and the invariant embryonic lineage may overcome the ectopic expression. Alternatively, it is possible that there are no somatic lineage defects because the critical MES-4 targeted germline genes that cause the L2 developmental delay and chemotaxis defect are not yet sufficiently misexpressed to cause these defects in the embryo. It is also possible

that the critical MES-4 targeted germline genes that are causing the defects are expressed, but not at a sufficient level to cause a defect. Regardless, the lack of somatic embryonic lineage defects raised the possibility that the L2 developmental delay and chemotaxis defects in *spr-5; met-2* mutants are not due to a failure to properly specify a specific lineage during embryogenesis.

Although we do not observe any somatic defects in the embryonic lineage of *spr-5; met-2* mutants, we detect a significant delay in duration of the germline blastomere P_4 before dividing to give rise to the primordial germ cells Z2 and Z3 in *spr-5; met-2* mutants compared to Wild Type. This delay occurs despite there being no corresponding delay in the P_4 sister cell D, arguing that the P_4 delay is highly specific to the germline lineage. While it remains unclear why there is a delay in the P_4 cell division, our single-cell data demonstrate that MES-4 targeted genes fail to be fully expressed in the germline cluster (7). Therefore, it is possible that the failure to activate the transcription of the MES-4 germline genes causes a delay in the P_4 cell division. Previously, it has been shown that the target of SPR-5, H3K4me2, is specifically lost during the P_4 cell division to generate Z2 and Z3. Thus, it is possible that an increase in H3K4me2 in *spr-5; met-2* results in a failure to lose H3K4me2 in Z2 and Z3. This in turn could lead to the delay in the P_4 cell division, but this possibility remains to be examined.

The lack of somatic embryonic lineage defects in *spr-5; met-2* mutants raised the possibility that the chemotaxis defect is due to the ongoing ectopic expression of MES-4 targeted germline genes. If this were the case, we would expect that eliminating the ectopic expression in worms that already have a chemotaxis defect would eliminate the chemotaxis defect. To test this possibility, we performed *mes-4* RNAi beginning at the

L2 stage in *spr-5; met-2* mutants to block the ectopic expression of MES-4 targeted germline genes, when these mutants already have a chemotaxis defect. Importantly, at the L2 larval stage *C. elegans* already have a completely intact nervous system. Remarkably, we find that knock down of *mes-4* beginning at the L2 stage significantly rescues normal chemotaxis behavior in the same animals that were previously defective at the L2 stage. MES-4 is thought to function in propagating the inappropriate expression of MES-4 targeted germline genes in somatic tissues and may not be completely required to maintain the ectopic expression of these genes. Therefore, we also tried to block the ectopic expression of germline genes by performing RNA interference against the germline transcription factor LSL-1. LSL-1 is required for the transcription of many germline genes, including many MES-4 targeted germline genes. Knock down of *lsl-1* beginning at the L2 stage in *spr-5; met-2* mutants resulted in an even stronger rescue of normal chemotaxis behavior than knock down of *mes-4.* Taken together these results suggest that the ectopic expression of MES-4 targeted germline genes actively interferes with normal chemotaxis behavior in a completely intact nervous system.

It is unclear how the ectopic expression of germline genes actively interferes with normal chemotaxis behavior. One possibility is that some normal component of neuronal function is blocked at the transcription level. For example the expression of synaptic proteins or a chemoreceptor could be actively inhibited in *spr-5; met-2* mutants. An alternative possibility that is not mutually exclusive is that some aspect of germline function actively interferes with neuronal function. For example, meiosis genes could cause inappropriate chromosomal condensation. Regardless, our finding that ectopic

transcription actively interferes with an intact nervous system has implications for neurodevelopmental diseases, such as Kabuki Syndrome and LSD1 patients. Based on our results, it is possible that the intellectual disability or altered behavior in these patients could be due to an ongoing defect in a properly formed nervous system. Even though these patients have *de novo* mutations that were inherited from one of their parents or mutated in the very early embryo, it is possible that the inappropriately inherited chromatin results in a defect that only manifests itself later in development because the transcription factors that are required to interact with the permissive chromatin are only activated in certain differentiated cell types, such as neurons. Regardless of the etiology, if the nervous system defects in these patients are due to the ongoing ectopic expression of genes, it may be possible to rescue these defects by turning off the ectopic transcription. Consistent with the possibility, it has been recently shown that behavioral defects in the MeCP2 Rhett Syndrome mouse model can be rescued by re-expressing MeCP2 in adult mice[27-29]. Our data provide a possible explanation for this result.

MATERIALS AND METHODS

Strains

All *C. elegans* strains were cultured at 20˚C on 60 mm nematode growth media (NGM) agar plates with OP50 bacteria grown in Luria Broth (LB). Strains used were: N2: wild-type (Bristol isolate); The *C. elegans spr-5 (by101)(I)* strain was provided by R. Baumeister (Albert Ludwig University of Freiburg, Germany); The MT13293: *met-2 (n4256)(III) strain* was provided by R. Horvitz (Massachusetts Institute of Technology, MA, USA); the JIM113: ujIs113 [Ppie-1::H2B::mCherry, unc-119(+); Pnhr-2::HIS-24::mCherry, unc-119(+) II, the strain was provided by J. Murray (University of Pennsylvania, PA, USA). *spr-5(b101)/tmC27[unc-75(tmls1239)]* I*; met-2(n4256)]/qC1 qIs26 [lag-2::GFP + pRF4 rol-6(su1006)]* III strain was created to maintain as double heterozygote. Automated cell lineage strain *spr-5(b101)/tmC27[unc-75(tmls1239)]* I; JIM113: ujIs113 [Ppie-1::H2B::mCherry, unc-119(+); Pnhr-2::HIS-24::mCherry, unc-119(+) II; *met-2(n4256)]/qC1 qIs26 [lag-2::GFP + pRF4 rol-6(su1006)]* III.

Single-worm genotyping

Single animals were picked into 5–10 ml of lysis buffer (50 mM KCl, 10 mM Tris-HCl (pH 8.3), 2.5 mM MgCl2, 0.45% NP-40, 0.45% Tween-20, 0.01% gelatin) and incubated at 65˚C for 1 hr followed by 95˚C for 30 min. PCR reactions were performed with AmpliTaq Gold (Invitrogen) according to the manufacturer's protocol and reactions were resolved on agarose gels. The following genotyping primers were used: spr-5 (*by101*)I: fwd: AAACACGTGGCTCCATGAAT, rev(wt):GAGGTTTTGAGGGGTTCCAT, rev(mut):CTTGAAACAGACTTGAACATCAAAGATCGG; met-2 (n4256):

fwd(wt):GTCACATCACCTGCATCAGC, rev(wt):ATTTCATTACGGC TGCCAAC, fwd(mut):ATTCGAAAAATGGACCGTTG, rev(mut):TCTATTCCCAGGAGCCAATG;

Quantitative PCR

Total RNA was isolated using TRIzol reagent (Invitrogen) from synchronized embryos at room temperature (21°C-22°C). cDNA synthesis and qPCR were carried out as previously described[11]. mRNA was quantified by real-time PCR, using iO SYBR Green Supermix (BioRad). The following primers were used: *htp-1* (ATTCGGAGGACAGTGACACAA and GTGCTTTCTCGAGAGACTCAGTTATATC) *cpb-1* (GTGCTGATTGATTGGCCTCG and CCGTTACAGCGCGTGAACCG); *rmh-1* (TGTAGTCATTATGCCAAGTATCTGC and ATCTGTTACTCGTATCTGTAGTAGCC); *ftr-1* (TCCGCTCACTTCGAATACGG and TACCATCGCGATTGTGAGC); *fbxa-101* (TATCGAAGACAAGCTCGCCG and TGCGAACGGAAATCCAATCG); *ama*-1 (TACCTACACTCCAAGTCCATCG and CGATGTTGGAGAGTACTGAG). Each mRNA expression was normalized to *ama-1* control. Fold enrichment was calculated mutant/N2.

RNA interference

Escherichia coli HT115 transformed with a vector expressing dsRNA of *mes-4*, *lsl-1*, L4440 and *ama-1* was obtained from the Ahringer library (Source BioScience). RNAi bacteria and empty vector control were grown at 37C and seeded on RNAi plates (standard NGM plates containing ampicillin @100 mg/ml and isopropylthiogalactoside (IPTG @0.4 mM)). RNAi plates were left at room temperature to induce for at least 24 hr.

Chemotaxis assay

The experimental 60mm NGM plate was divided into four quadrants. 5μL of control LB/vehicle was pipetted into two quadrants. In the other two quadrants an experimental sample, 5μL of *Escherichia coli* OP50 was pipetted. The plate was placed into the fume hood for drying purposes. Worms were collected and rinsed 3 times with M9 buffer (22 mM KH_2PO_4, 42 mM Na_2HPO_4, 86 mM NaCl, and 1 mM $MgSO_4$). A total of 50 worms per plate were placed into the center of the experimental plate. The plates were then incubated at 20°C for 1-3 hours. Worms that moved to any of the quadrants after the incubation time was recorded. The Chemotaxis Index was calculated as follows; (E1 + E2) – (C1 + C2)/ Total worms moved.

Lineage tracing

The experiment was carried out as previously described[26] with some minor changes. While using the 20μm beads we noticed that the embryo was moving away from the focus. We decided to use an NGM plate as a platform to place the embryo. The alternative approach is described in detail[30].

Single-cell sequencing and data analysis

Single-cell isolation was performed according to Packer et al.[31] with minor modifications. For the 100-200 cell stage, only one enzyme (Chitinase) was used for embryo eggshell disruption and single cells dissociation[30]. Single-cell RNA sequencing was performed using the 10 x Genomics single-cell capturing system. Single cells were isolated from approximately 1,000 worms per strain. Ranging from 1,000 to 3,000 cells, were loaded on the 10 X Genomics Chromium Controller. Single-cell cDNA libraries were prepared using the Chromium Next GEM Single Cell 3' LT Reagent Kits v3.1 (Dual Index).

Libraries were sequenced by FSU College of Medicine Translational Science Laboratory (Florida State University, FL, USA). After the 10x QC, N2 has 219 estimated number of Cells, 62,491 mean Reads per Cell and 1,067 Median Genes per cell. *spr-5; met-2* has 686 estimated Number of Cells, 20,763 mean Reads per Cell and526 median genes per cell. Sequencing: 150 bp paired-end reads (Illumina NextSeq 600), samples received 28 million paired-end reads and Cell Ranger Software Suite 3.0.2 (10x Genomics) was used for the alignment of the single-cell RNA-seq output reads and generation of feature, barcode and matrices. Seurat analysis was used for cluster formation.

ACKNOWLEDGEMENTS

We are grateful to members of the Katz lab for their helpful discussion; J. Murray, C. Huynh, and A. Santella for advice on lineage tracing and single cell RNAseq. We thank J. Murray for the JIM113 strain as well as R. Horvitz and the *Caenorhabditis* Genetics Center (funded by NIH P40 OD010440) for additional strains. Juan D. Rodriguez was supported by NIH National Institute of Child Health and Human Development 3F31HD100145-03S1 grant and the training grant (T32GM008490-21). This work was funded by a grant to D.J.K. (NSF IOS1354998).

FIGURES and FIGURES LEGENDS

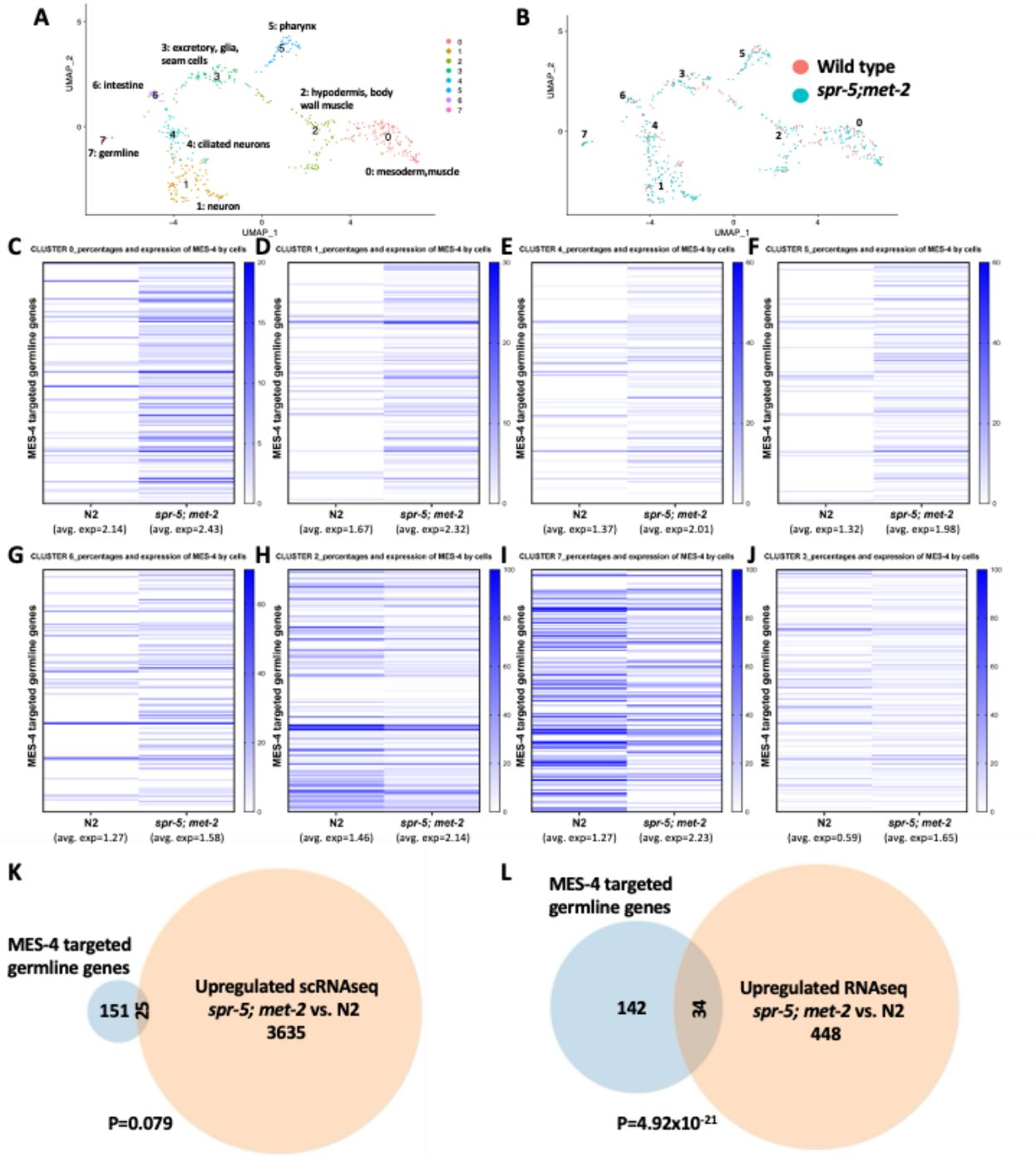
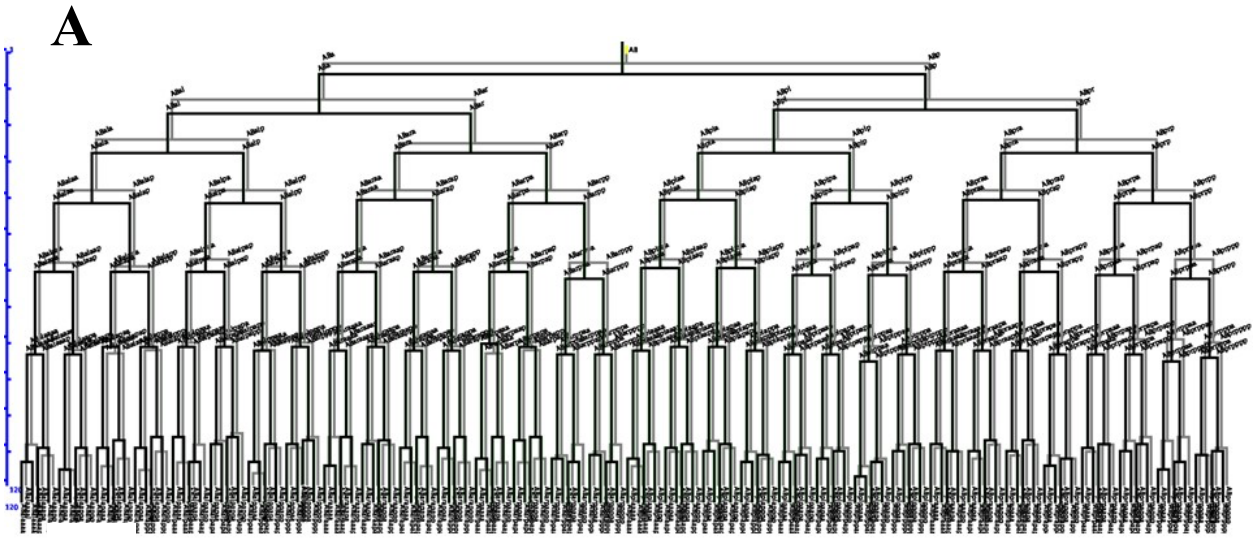
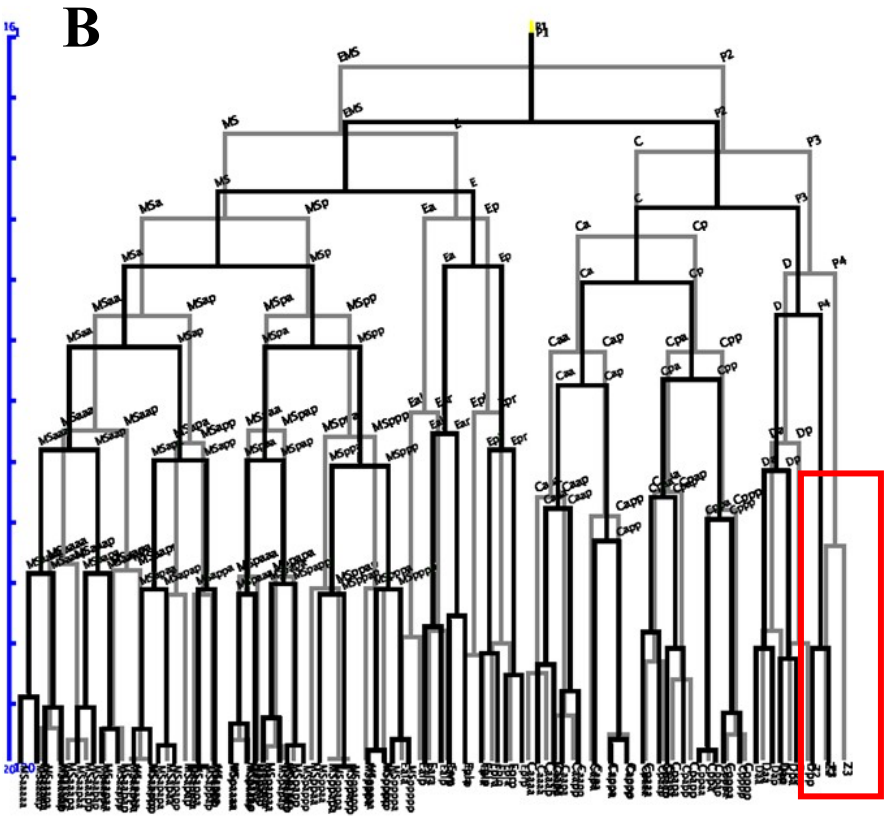


Figure 1. MES-4 targeted germline genes begin to be ectopically expressed during embryogenesis in *spr-5; met-2* mutants. (A) UMAP projection of all 219 Wild Type cells and 686 *spr-5; met-2* mutant cells from single-cell RNAseq formed 8 clusters. (B) All Wild Type (red) and *spr-5; met-2* double mutant (green) cells fall within these 8 clusters. (C-J) Heat maps from all 8 clusters (from A) showing the percentage of cells that

express each of the 197 MES-4 targeted germline genes in Wild Type (N2) compared to *spr-5; met-2* mutants. The average gene expression across all 176 MES-4 targeted germline genes for each cluster is shown below the cluster. (K,L) Overlap between MES-4 targeted germline genes and total genes significantly upregulated in *spr-5; met-2* mutants in any of the 8 individual single-cell RNAseq clusters (K), or significantly upregulated in *spr-5; met-2* mutants at the L1 larval stage (L) from Carpenter et al. 2021. Significance in K and L was determined by a hypergeometric test.


A


B

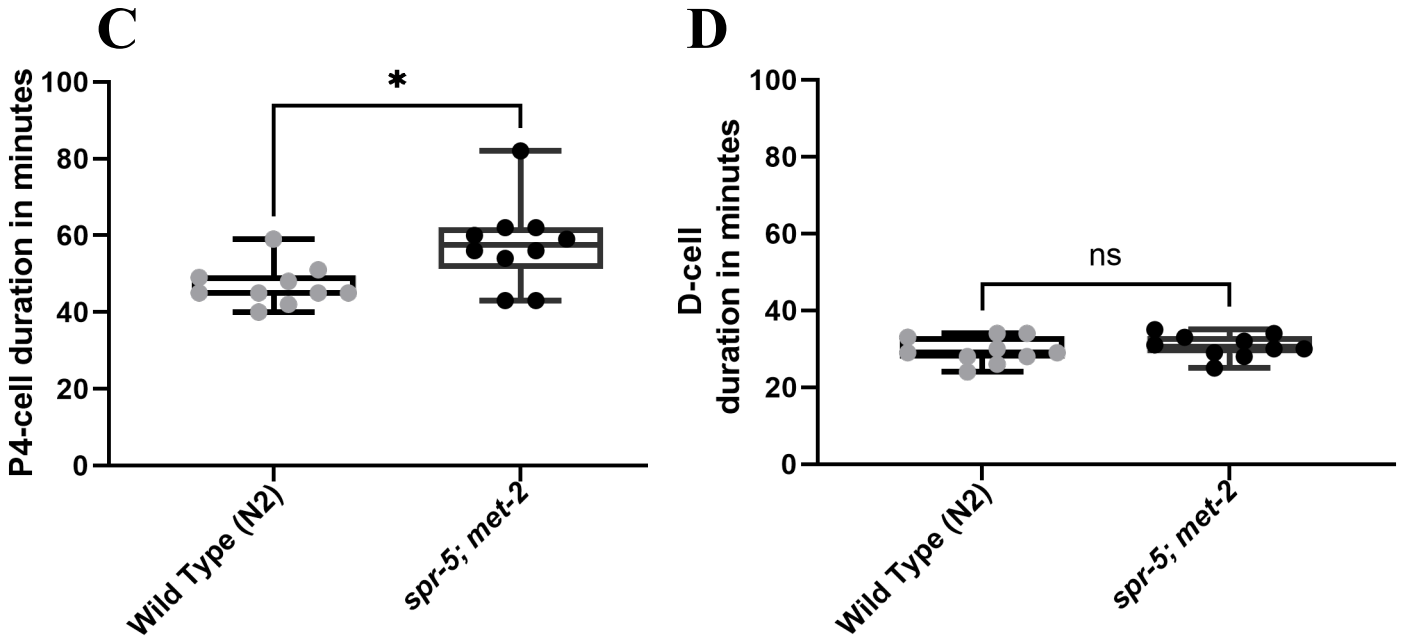


Figure 2. Comparison of the embryonic cell lineage in *spr-5; met-2* versus Wild Type (N2). Overlap between *spr-5; met-2* (black) and Wild Type (grey) in the AB (A) and P lineage (B) from the 2-cell stage through the ~200-cell stage. A total of 5 lineages of each were analyzed. The Y-axis indicates the end of cell division in minutes. The red box highlights the only observed defect, a delay in the division of the germline blastomere P_4 to the primordial germ cells, Z2 and Z3, quantified by duration in minutes in (C). (D) The corresponding duration in minutes of the cell D, immediately adjacent to P_4 in the embryo. (C,D) N= 10 for Wild Type and N=9 for *spr-5; met-2.* The error bars in C and D are S.E.M. Significance in C and D was determined by an unpaired t-test. *≤0.05

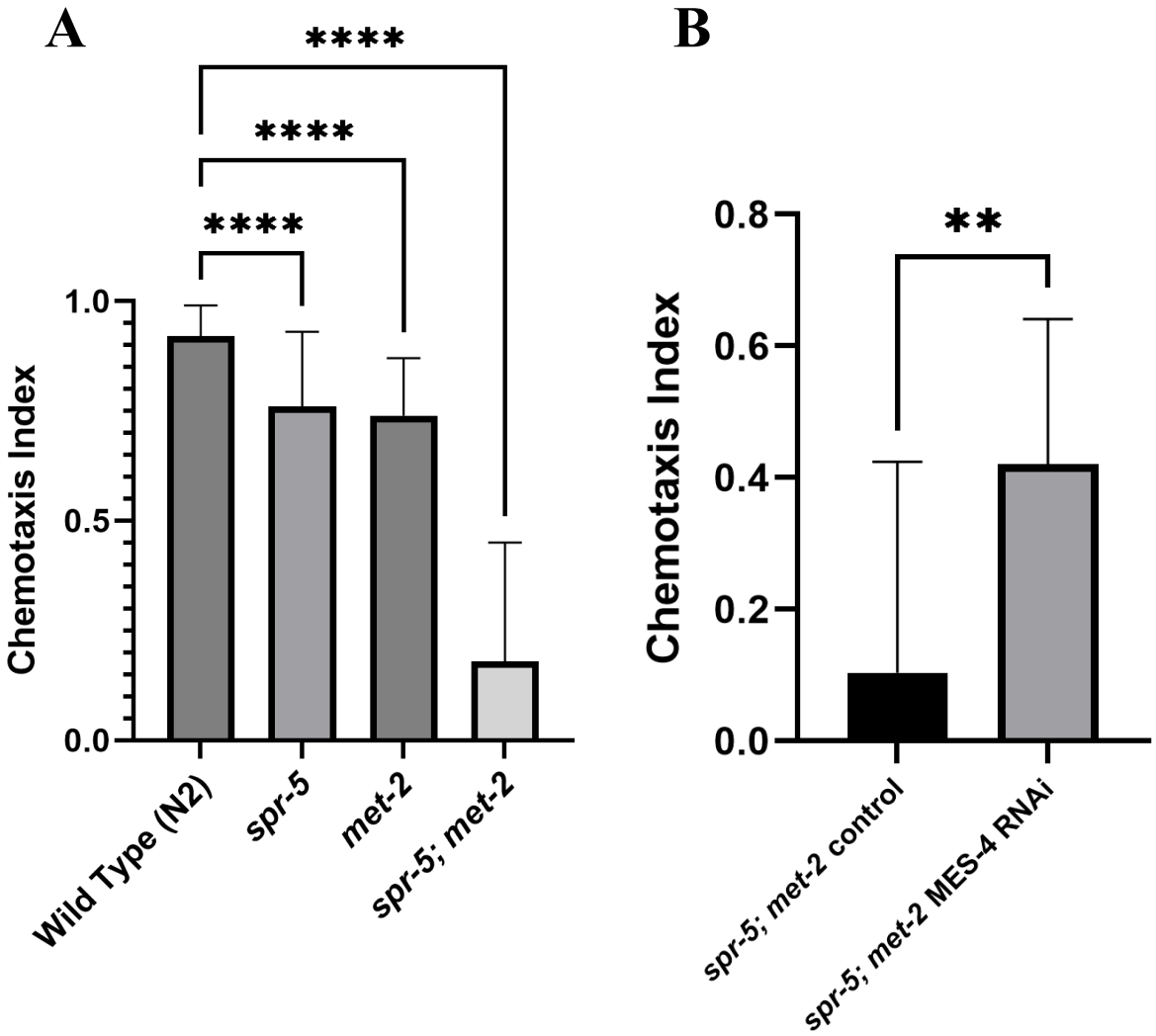


Figure 3. Chemotaxis defect in *spr-5; met-2* progeny. (A) The chemotaxis index of Wild Type (N2) N=1008, *spr-5* N=1661, *met-2* N=1117 and *spr-5; met-2* N=263. (B) The chemotaxis index of *spr-5; met-2* mutants is rescued by *mes-4* RNAi (N= 948) compared to control (L4440) RNAi (N=690). The error bars represent the S.E.D.. Significance was calculated in (A) by one-way ANOVA and (B) by unpaired t-test. ****≤0.0001, **≤0.01.

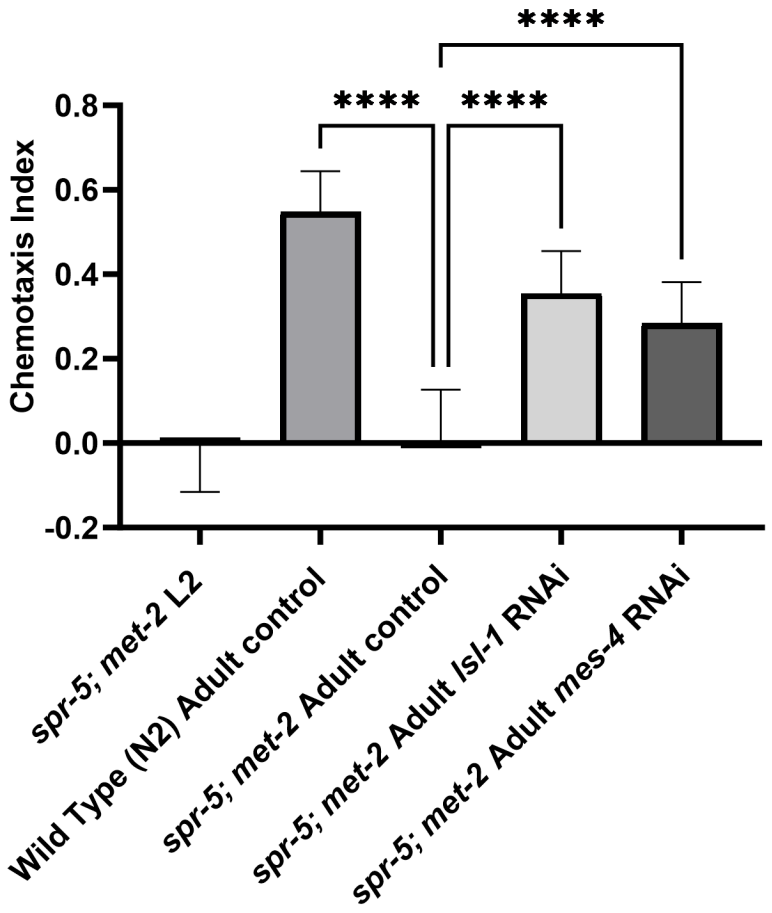


Figure 4. Blocking the ectopic transcription of MES-4 targeted germline genes in L2 *spr-5; met-2* mutants that already have a chemotaxis defect restores normal chemotaxis. The chemotaxis index of L2 *spr-5; met-2* double mutants (N=2950) is significantly rescued by *mes-4* (N=763) and *lsl-1* (776) RNAi in the same worms that previously had a chemotaxis defect at the L2 stage to levels that are similar to Wild Type (N2) (N=1003) under these conditions. In contrast, control (L4440) RNAi does not rescue (N=603). The error bars represent the S.E.M.. Significance was calculated by an unpaired t-test. ****= <0.0001

References Cited

1Adam, M. P., Banka, S., Bjornsson, H. T., Bodamer, O., Chudley, A. E., Harris, J., Kawame, H., Lanpher, B. C., Lindsley, A. W., Merla, G., Miyake, N., Okamoto, N., Stumpel, C. T., Niikawa, N. & Kabuki Syndrome Medical Advisory, B. Kabuki syndrome: international consensus diagnostic criteria. *J Med Genet* 56, 89-95, doi:10.1136/jmedgenet-2018-105625 (2019).
2Chong, J. X., Yu, J. H., Lorentzen, P., Park, K. M., Jamal, S. M., Tabor, H. K., Rauch, A., Saenz, M. S., Boltshauser, E., Patterson, K. E., Nickerson, D. A. & Bamshad, M. J. Gene discovery for Mendelian conditions via social networking: de novo variants in KDM1A cause developmental delay and distinctive facial features. *Genet Med* 18, 788-795, doi:10.1038/gim.2015.161 (2016).
3Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S. & Greenblatt, J. Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* 23, 4207-4218 (2003).
4Cenik, B. K. & Shilatifard, A. COMPASS and SWI/SNF complexes in development and disease. *Nat Rev Genet* 22, 38-58, doi:10.1038/s41576-020-0278-0 (2021).
5Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* 128, 707-719 (2007).
6Lee, T. W. & Katz, D. J. Hansel, Gretel, and the Consequences of Failing to Remove Histone Methylation Breadcrumbs. *Trends Genet* 36, 160-176, doi:10.1016/j.tig.2019.12.004 (2020).
7Katz, D. J., Edwards, T. M., Reinke, V. & Kelly, W. G. A C. elegans LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory. *Cell* 137, 308-320 (2009).
8Wasson, J. A., Simon, A. K., Myrick, D. A., Wolf, G., Driscoll, S., Pfaff, S. L., Macfarlan, T. S. & Katz, D. J. Maternally provided LSD1/KDM1A enables the maternal-to-zygotic transition and prevents defects that manifest postnatally. *Elife* 5, doi:10.7554/eLife.08848 (2016).
9Ancelin, K., Syx, L., Borensztein, M., Ranisavljevic, N., Vassilev, I., Briseno-Roa, L., Liu, T., Metzger, E., Servant, N., Barillot, E., Chen, C. J., Schule, R. & Heard, E. Maternal LSD1/KDM1A is an essential regulator of chromatin and transcription landscapes during zygotic genome activation. *Elife* 5, doi:10.7554/eLife.08851 (2016).
10Padeken, J., Methot, S. P. & Gasser, S. M. Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nat Rev Mol Cell Biol* 23, 623-640, doi:10.1038/s41580-022-00483-w (2022).
11Kerr, S. C., Ruppersburg, C. C., Francis, J. W. & Katz, D. J. SPR-5 and MET-2 function cooperatively to reestablish an epigenetic ground state during passage through the germ line. *Proc Natl Acad Sci U S A* 111, 9509-9514, doi:10.1073/pnas.1321843111 1321843111 [pii] (2014).
12Greer, E. L., Beese-Sims, S. E., Brookes, E., Spadafora, R., Zhu, Y., Rothbart, S. B., Aristizabal-Corrales, D., Chen, S., Badeaux, A. I., Jin, Q., Wang, W., Strahl, B. D., Colaiacovo, M. P. & Shi, Y. A Histone Methylation Network Regulates Transgenerational Epigenetic Memory in C. elegans. *Cell Rep*, doi:S2211-1247(14)00158-2 [pii] 10.1016/j.celrep.2014.02.044 (2014).
13Andersen, E. C. & Horvitz, H. R. Two C. elegans histone methyltransferases repress lin-3 EGF transcription to inhibit vulval development. *Development* 134, 2991-2999 (2007).

14Kim, J., Zhao, H., Dan, J., Kim, S., Hardikar, S., Hollowell, D., Lin, K., Lu, Y., Takata, Y., Shen, J. & Chen, T. Maternal Setdb1 Is Required for Meiotic Progression and Preimplantation Development in Mouse. *PLoS Genet* 12, e1005970, doi:10.1371/journal.pgen.1005970 (2016).
15Rechtsteiner, A., Ercan, S., Takasaki, T., Phippen, T. M., Egelhofer, T. A., Wang, W., Kimura, H., Lieb, J. D. & Strome, S. The histone H3K36 methyltransferase MES-4 acts epigenetically to transmit the memory of germline gene expression to progeny. *PLoS Genet* 6, doi:10.1371/journal.pgen.1001091 (2010).
16Furuhashi, H., Takasaki, T., Rechtsteiner, A., Li, T., Kimura, H., Checchi, P. M., Strome, S. & Kelly, W. G. Trans-generational epigenetic regulation of C. elegans primordial germ cells. *Epigenetics Chromatin* 3, 15, doi:1756-8935-3-15 [pii] 10.1186/1756-8935-3-15 (2010).
17Rodriguez-Crespo, D., Nanchen, M., Rajopadhye, S. & Wicky, C. The zinc-finger transcription factor LSL-1 is a major regulator of the germline transcriptional program in Caenorhabditis elegans. *Genetics* 221, doi:10.1093/genetics/iyac039 (2022).
18Carpenter, B. S., Lee, T. W., Plott, C. F., Rodriguez, J. D., Brockett, J. S., Myrick, D. A. & Katz, D. J. Caenorhabditis elegans establishes germline versus soma by balancing inherited histone methylation. *Development* 148, doi:10.1242/dev.196600 (2021).
19Unhavaithaya, Y., Shin, T. H., Miliaras, N., Lee, J., Oyama, T. & Mello, C. C. MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in C. elegans. *Cell* 111, 991-1002 (2002).
20Petrella, L. N., Wang, W., Spike, C. A., Rechtsteiner, A., Reinke, V. & Strome, S. synMuv B proteins antagonize germline fate in the intestine and ensure C. elegans survival. *Development* 138, 1069-1079, doi:10.1242/dev.059501 (2011).
21Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev Biol* 100, 64-119 (1983).
22Sulston, J. E. & Horvitz, H. R. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev Biol* 56, 110-156, doi:10.1016/0012-1606(77)90158-0 (1977).
23Bao, Z., Murray, J. I., Boyle, T., Ooi, S. L., Sandel, M. J. & Waterston, R. H. Automated cell lineage tracing in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 103, 2707-2712, doi:10.1073/pnas.0511111103 (2006).
24Robertson, S. M., Shetty, P. & Lin, R. Identification of lineage-specific zygotic transcripts in early Caenorhabditis elegans embryos. *Dev Biol* 276, 493-507, doi:10.1016/j.ydbio.2004.09.015 (2004).
25Baugh, L. R., Hill, A. A., Slonim, D. K., Brown, E. L. & Hunter, C. P. Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. *Development* 130, 889-900, doi:10.1242/dev.00302 (2003).
26Boyle, T. J., Bao, Z., Murray, J. I., Araya, C. L. & Waterston, R. H. AceTree: a tool for visual analysis of Caenorhabditis elegans embryogenesis. *BMC Bioinformatics* 7, 275, doi:10.1186/1471-2105-7-275 (2006).
27Sztainberg, Y., Chen, H. M., Swann, J. W., Hao, S., Tang, B., Wu, Z., Tang, J., Wan, Y. W., Liu, Z., Rigo, F. & Zoghbi, H. Y. Reversal of phenotypes in MECP2 duplication mice using genetic rescue or antisense oligonucleotides. *Nature* 528, 123-126, doi:10.1038/nature16159 (2015).
28Giacometti, E., Luikenhuis, S., Beard, C. & Jaenisch, R. Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. *Proc Natl Acad Sci U S A* 104, 1931-1936, doi:10.1073/pnas.0610593104 (2007).

29Garg, S. K., Lioy, D. T., Cheval, H., McGann, J. C., Bissonnette, J. M., Murtha, M. J., Foust, K. D., Kaspar, B. K., Bird, A. & Mandel, G. Systemic delivery of MeCP2 rescues behavioral and cellular deficits in female mouse models of Rett syndrome. *J Neurosci* 33, 13612-13620, doi:10.1523/JNEUROSCI.1854-13.2013 (2013).
30Rodriguez, J. D. & Katz, D. J. Lineage Tracing and Single-Cell RNA-seq in C. elegans to Analyze Transgenerational Epigenetic Phenotypes Inherited from Germ Cells. *Methods Mol Biol* 2677, 61-79, doi:10.1007/978-1-0716-3259-8_3 (2023).
31Packer, J. S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., Stefanik, D., Tan, K., Trapnell, C., Kim, J., Waterston, R. H. & Murray, J. I. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. *Science* 365, doi:10.1126/science.aax1971 (2019).