

Induction of kidney-related gene programs through co-option of *SALL1* in mole ovotestes

Magdalena Schindler^{1,2}, Marco Osterwalder³⁻⁵, Izabela Harabula⁶, Lars Wittler⁷, Athanasia C. Tzika⁸, Dina K.N. Dechmann^{9,10}, Martin Vingron¹¹, Axel Visel^{5,12,13}, Stefan Haas¹¹ and Francisca M. Real^{1,2,*}

¹Gene Regulation & Evolution. Max Planck Institute for Molecular Genetics, Berlin, Germany

²Institute for Medical and Human Genetics, Charité - Universitätsmedizin Berlin, Germany

³Department for BioMedical Research (DBMR), University of Bern, Bern, Switzerland.

⁴Department of Cardiology, Bern University Hospital, Bern, Switzerland

⁵Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA

⁶Epigenetic Regulation and Chromatin Architecture. Max-Delbrück-Centrum für Molekulare Medizin (MDC) Berlin, Germany

⁷Department of Developmental Genetics, Transgenic Unit. Max Planck Institute for Molecular Genetics, Berlin, Germany

⁸Department of Genetics & Evolution. University of Geneva, Switzerland

⁹Department of Migration, Max Planck Institute for Animal Behavior, Radolfzell, Germany

¹⁰Department of Biology, University of Konstanz, Konstanz, Germany

¹¹Department of Computational Molecular Biology, Max Planck Institute for Molecular Genetics, Berlin, Germany

¹²Department of Energy Joint Genome Institute, Berkeley, California 94720, USA

¹³School of Natural Sciences, University of California, Merced, California 95343, USA

* Author for correspondence: martinez@molgen.mpg.de

ABSTRACT

Changes in gene expression represent an important source of phenotypic innovation. Yet, how such changes emerge and impact the evolution of traits remains elusive. Here, we explore the molecular mechanisms associated with the development of masculinizing ovotestes in female moles. By performing integrative analyses of epigenetic and transcriptional data in mole and mouse, we identified the co-option of *SALL1* expression in mole ovotestes formation. Chromosome conformation capture analyses highlight a striking conservation of the 3D organization at the *SALL1* locus, but an evolutionary divergence of enhancer activity. Interspecies reporter assays support the capability of mole-specific enhancers to activate

transcription in urogenital tissues. Through overexpression experiments in transgenic mice, we further demonstrate the capability of *SALL1* to induce kidney-related gene programs, which are a signature of mole ovotestes. Our results highlight the co-option of gene expression, through changes in enhancer activity, as a plausible mechanism for the evolution of traits.

INTRODUCTION

Coordinated gene expression represents the cornerstone of developmental processes and homeostasis. In animals, transcription is mainly controlled by the action of cis-regulatory elements (CREs), such as enhancers, which control gene expression patterns with spatial and temporal precision. CREs control tissue-specific aspects of gene expression, acting in cooperation to constitute complex and pleiotropic gene expression patterns¹. To exert their function, CREs enter into physical proximity with gene promoters, mediated by the 3D folding of chromatin. CRE-promoter interactions are framed within topologically associating domains (TADs), which are 3D chromatin structures containing loci that interact with increased frequencies, and shielded from the regulatory influence of other genomic regions^{2,3}.

Coding mutations generally alter all the different functions of a gene, thus inducing systemic effects that might be detrimental for the development of an organism. In contrast, mutations in CREs display tissue-specific effects, thus preserving essential gene functions in other tissues. Consistently, the multiplicity of CREs can confer variations in expression patterns that contribute to gene pleiotropy, and support the rapid evolvability of these non-coding elements⁴. Indeed, mutations altering regulatory elements have been associated with the emergence of certain traits, such as the evolution of limbs in ungulate animals⁵.

Furthermore, the repurposing of a gene or regulatory element to a new function through a co-option process also represents an important source for phenotypic innovation⁶⁻⁸. This mechanism has been exemplified in the evolution of the neural crest cells in vertebrates through the acquisition of new regulatory elements for the *SoxE* family genes⁹. Another relevant example of co-option is illustrated on how the propagation of retroviruses in the mammalian genomes has shaped the regulatory landscape of the immune system¹⁰. Therefore, variations in gene expression and function, through CRE mutations, underlie the evolution of certain phenotypic traits and can represent the basis for species adaptation.

A prominent example of phenotypic evolution is observed in Talpid moles. Female moles consistently develop ovotestes, instead of ovaries like most mammalian species. These gonads are composed of ovarian tissue, supporting a fertile function, and a sterile testicular portion that secretes male hormones. These hormones exert a masculinizing effect in female

moles, increasing muscle strength and aggression, aspects that likely contribute to their adaptation to subterranean environments. In a previous study, we demonstrated that the evolution of ovotestes is associated with the reorganization of TADs, which alter CRE-promoter interactions and gene expression patterns¹¹. In particular, a large inversion relocates active enhancers in the vicinity of the pro-testicular gene *FGF9*, whose ectopic expression in female gonads leads to meiosis inhibition and masculinization. In addition, a duplication of enhancer elements is associated with the increased expression of *CYP17A1* encoding an enzyme for male hormone synthesis and increase of muscle strength. While the observed regulatory changes at these loci partially explain the mole phenotype, it is plausible that additional mechanisms contribute to the evolution of this trait.

In this study, we further investigate the molecular mechanisms associated with mole ovotestis development. Using integrative epigenetic and transcriptional approaches in mole and mouse, we identify that the expression of the transcription factor *SALL1* has been co-opted in the formation of XX testicular tissue in the Iberian mole *Talpa occidentalis*, through changes in CREs. Our finding is further supported by expression analyses in closely related species developing normal ovaries, like shrews and hedgehogs. We determine the regulatory landscape of this gene, highlighting an evolutionary conserved TAD structure, but with divergent enhancer activity. Through *in vivo* interspecies reporter assays, we reveal the potential of enhancer elements to evolve new activity domains in moles. By using transgenic mice that overexpress *Sall1* in ovaries, we demonstrate the capacity of this factor to activate kidney-related gene programs that are also observed during mole ovotestis formation. Altogether, our results further extend our understanding of the molecular basis of a unique trait, highlighting the important role of regulatory variation in evolution.

RESULTS

Evolutionary conservation of mammalian gonadal enhancers

CREs represent a major source of tissue-specific gene expression¹. We previously explored the regulatory landscape of mole developing gonads, at an early postnatal stage (7 days *post-partum* – stage P7)¹¹. At this developmental time-point, testicular and ovarian tissues from female ovotestes are first morphologically discernable and can be microdissected (**Figure 1A**). Furthermore, Leydig cells of the testicular part differentiate and produce testosterone, whereas meiosis initiates in the ovarian part, an event considered as one of the earliest signs of female gonadogenesis in mammals¹². We identified regulatory elements in mole gonads by performing ChIP-seq experiments against a combination of histone marks, H3K27ac together with H3K4me1 and H3K4me3, for the distinction of enhancers and promoters, respectively. By

using the tool CRUP¹³, we combined these datasets in each sampled tissue to call and rank active regulatory regions according to their enhancer probability score (**Figure 1B**).

To explore the degree of conservation of the enhancer landscape in moles, we generated analogous datasets from mouse gonads, at a time point when Leydig cells differentiate, and meiosis takes place (E13.5; **Figure 1B**). By comparing mole and mouse gonadal epigenetic datasets, we observed that from the 70,618 predicted enhancers in mole gonads approximately 65% are conserved to some extent at the sequence level, meaning they can be lifted over to the mouse genome (**Figure 1C**). However, only 25% of those enhancers are active in both species, meaning that they share an active enhancer signature in both mole and mouse gonads. Accordingly, approximately 40% of the predicted sequence conserved enhancers represent mole-specific regulatory regions and are thus potentially associated with characteristics of this species. Therefore, our results imply a repurposing of enhancer function during gonad evolution.

Co-option of *SALL1* expression in mole ovotestis formation

Our approach identified a subset of 6,419 mole-specific enhancers that are only active in the testicular part of the ovotestis and could potentially contribute to the development of this unique tissue. We then explored if these enhancers are associated with the acquisition of specific transcriptional signatures using RNA-seq datasets from the same developmental stage. We therefore jointly ranked enhancers by specificity in enhancer probability in the testicular part of the ovotestes and the specific expression of their putative target gene in the same tissue. We defined the putative target genes of each enhancer as the gene with the closest transcription start site to the enhancer region within the same TAD. This approach prioritizes genes whose respective regulatory domain contains enhancer elements specifically active in the testicular part compared to the ovary part and the male testis (**Figure 2A, Supplementary Table 1**). The top-ranking genes identified by this approach were *NPY* and *SALL1*. *NPY* is a hormone neuropeptide expressed in Leydig cells^{14,15}, whereas *SALL1* is a transcription regulator involved in cell fate decision¹⁶. *SALL1* is usually expressed during development in embryonic tissues, including eye, neural tube, limb or kidney¹⁷. Strikingly, our RNA-seq data revealed that *SALL1* is highly expressed in the testicular part of mole ovotestes at P7, but not in the XY testis or the XX ovarian region. In fact, *SALL1* is highly expressed already in the early embryonic ovotestis and becomes specific to the testis part as the organ differentiates (**Figure 2B**). In humans, mutations in *SALL1* are associated with a congenital malformation syndrome affecting limbs, kidneys and ears (Townes Brocks syndrome, OMIM #107480)¹⁸ and misexpression has been linked to

certain types of androgen-producing ovarian tumors¹⁹, indicating that it might be involved in re-programming ovarian cells.

To further explore the spatio-temporal dynamics of *SALL1* expression, we performed immunostaining in mole gonads at different stages of development (**Figure 2C**). This analysis revealed that *SALL1* expression is specific to the mole female gonad, and importantly, this expression is spatially restricted to the medullary region of the developing ovotestis, which is the precursor of the testicular tissue. Double immunostaining for *SALL1* and *FOXL2*, a marker of female somatic cells²⁰, (**Figure 2D**) confirmed that *SALL1* expression is restricted to the testicular part of the ovotestis. This staining also revealed that cells that are simultaneously positive for *FOXL2* and *SALL1* form the spherules, which are equivalent to testis cords and considered "Sertoli-like" cells. Based on these results, the expression of *SALL1* detected in the RNA-seq from the ovarian part of the adult ovotestis (**Figure 2B**) is likely due to imperfect dissection of the tissue, which is especially challenging as the two gonadal compartments are intricately connected at this stage. Therefore, the expression pattern of *SALL1* is constant during the entire development and persists in adulthood, thus constituting this gene as a bona-fide marker for the testicular tissue of mole ovotestis.

We then explored the evolutionary conservation of *SALL1* expression in other mammalian species. We examined the pattern of expression of *Sall1* in mice by immunostaining and transcriptomic analyses. Immunostaining analyses showed a complete absence of *SALL1* protein in mouse gonads at embryonic stage E13.5, however, the protein could be detected in known *Sall1*-expressing tissues, such as the embryonic kidneys (**Supplementary Figure 1A**). This observation is extended to adulthood where RNA-seq data shows practically no expression in both males and females when compared to the mole (**Supplementary Figure 1B**). We further expanded our analysis of *SALL1* expression to also include species from the order *Eulipotyphla*, which are evolutionarily close to moles²¹. Specifically, we analyzed ovarian samples from the hedgehog *Atelerix albiventris*, as well as from the common shrew, *Sorex araneus*, the latter species belonging to the closest taxonomic group but developing normal ovaries. Immunostaining analyses showed the absence of *SALL1* expression in the gonads of these two species (**Figure 2E**). However, we could detect *SALL1* in other control tissues such as neural tube or kidney from hedgehog and shrew, proving the specificity of the antibody used (**Supplementary Figure 2A, B**). In addition, the absent of expression of *SALL1* in the ovaries of these species was further confirmed by RT-qPCR (**Supplementary Figure 2C, D**). Overall, these results indicate that *SALL1* expression has been acquired during the evolution of mole ovotestes.

Conserved 3D organization but divergent enhancers at the mole SALL1 locus

To define the regulatory landscape of *SALL1*, we examined previously published Hi-C data from different mole tissues¹¹ (**Figure 3A, Supplementary Figure 3**). Chromatin interaction maps revealed a large 1 Mb TAD, in which *SALL1* is the only protein-coding gene. The interaction profile of *SALL1* in the testicular part of the ovotestis was further explored at increased resolution through circular chromosome conformation capture (4C-seq), using the gene promoter as a viewpoint (**Figure 3B**). These experiments demonstrate prominent interactions of *SALL1* across the entire TAD, with a sharp decrease in contacts outside this domain. We then explored the degree of conservation of the *SALL1* interaction profile by comparing the mole against mouse data²². This comparison revealed that, despite notable differences in *SALL1* expression, the locus displays a remarkable preservation of its 3D structure across species (**Supplementary Figure 4A**).

Next, we overlaid the *SALL1* interaction profile to the epigenetic datasets, to identify potential regulatory elements (**Figure 3C**). This revealed several candidate enhancer regions that were active exclusively in the testicular part of the ovotestis. Specifically, we identified one putative enhancer element that is close to *SALL1* and unique for the testicular portion, as well as a distant cluster of four additional elements. This putative enhancer cluster is indeed in close physical proximity to the *SALL1* promoter, as denoted by a specific loop in the Hi-C map and an increase in contacts in the 4C profile. A zoom-in on these regions highlights the specificity of these enhancers for the testicular part of the ovotestes (**Figure 3D**). Consistently to its conserved 3D structure, these candidate enhancers lie in syntenic regions when aligned against mouse or shrew genomes (**Supplementary Figure 5**). However, a comparison with the respective mouse epigenetic datasets revealed that these elements were not active in mouse gonads (**Figure 3E**). Specific alignments of these five enhancers against mouse and shrew revealed only a partial degree of sequence conservation (**Supplementary Figure 6 and 7**).

To validate the activity of these putative enhancers *in vivo*, we tested the five mole regions for enhancer activity in mouse transgenic LacZ reporter assays²³ (E1-5; **Figure 3D**). Of note, these elements display active enhancer marks that are specific for the testis part of the ovotestes, but such marks are not present in mouse gonadal tissue. At E13.5, all regions tested showed reproducible tissue-restricted activity, thus confirming them as true enhancers (**Figure 4; Supplementary Figures 8-12**). Enhancer activity was observed in several tissues, such as the limbs or eyes, in which *Sall1* is known to be expressed. Interestingly, enhancer 3 displayed specific activity in kidneys, another *Sall1*-expressing tissue¹⁷, which is consistent with its predicted enhancer activity in mouse embryonic kidneys (**Supplementary Figure 4B**). While none of these enhancers induced reporter expression in developing gonads, enhancers 1, 2, 4

and 5 were active in the adjacent mesonephros. This tissue has the same ontogenetic origin as the gonads, and contributes to its cellular composition through cell migration²⁴. Furthermore, it has been previously shown that *SALL1* is expressed in the mesonephric duct of mice²⁵, a pattern that is also conserved in moles (**Supplementary Figure 13**).

We sought to investigate whether the divergence observed in the mole enhancer sequences included alterations in transcription factor binding sites that may explain the specific activation of *SALL1* in the ovotestis. To this end, we conducted a transcription factor enrichment analysis on the five mole enhancer sequences and compared them to mouse and shrew sequences. We further filtered putative motifs for transcription factors expressed in mole gonadal tissues. The results revealed a distinctive binding pattern among species (**Supplementary Table 2**), with minimal overlap in the most significant transcription factor bindings (**Supplementary Figure 14A**). Moreover, we observed higher expression of some top-ranked transcription factors, such as *IRF4* or *FOXP1*, in the testicular part of the ovotestes compared to mice (**Supplementary Figure 14B**), which could account for the lack of LacZ activity in mouse gonads. This observation, together with the moderate sequence conservation compared to other mammals suggests that the evolution of enhancers in the regulatory domain of *SALL1* may have driven its expression in the testicular part of mole ovotestis.

***SALL1* expression triggers kidney-related gene programs during ovarian development**

To investigate the effects of *Sall1* expression during early gonadal development, we induced its expression in the mouse ovary. For this purpose, we created a BAC construct to overexpress *Sall1* in somatic ovarian cells (**Figure 5A**). The BAC contains the regulatory elements and the promoter of the *Wt1* gene, which is constitutively expressed in gonadal somatic cells²⁶, but the gene is replaced by the coding sequence of *Sall1*. Through *PiggyBac* transgenesis, we integrated this construct into female mouse embryonic stem cells (mESC), which were subsequently used to generate transgenic mice through morula aggregation. In contrast to wildtype controls, mutant ovaries express *Sall1* in somatic cells, as denoted by the overlapping signal with *Foxl2*, a bona-fide marker for female somatic ovarian cells²⁰ (**Figure 5B**). However, at the phenotypic level, adult female mice did not show major morphological gonadal alterations and bred normally. Similarly, *Sall1*-overexpressing males develop normal testes and did not show any sign of reduced fertility (**Figure Supplementary 15**). This suggests that *Sall1*, by itself, is not sufficient to induce the development of testicular structures, nor to disrupt normal testis development.

To gain further insights into the molecular signatures of *Sall1* ovarian expression, we performed RNA-seq in gonads from mutants and litter-mate controls at E13.5. This analysis revealed around 400 deregulated genes where *Sall1* is the most significantly up-regulated gene (**Figure 5C, Supplementary Table 3**). To understand the consequences of *Sall1* expression in female gonads, we compared the deregulated genes in the mutant ovaries to those specifically expressed in the testicular part of the ovotestis. We found 56 upregulated and 36 downregulated genes that are shared between the mutant mouse gonad and the testicular part of the mole ovotestis. Gene ontology analyses revealed no significant enrichment for the downregulated genes. However, the upregulated genes were enriched in terms related to the development of the kidney, a tissue in which *SALL1* is consistently expressed across mammalian species, as well as to ureteric bud morphogenesis and mesonephros development (**Figure 5D, Supplementary Figure 16**).

The migration of somatic cells from mesonephros to gonad is a characteristic process of testis development, not occurring in ovaries^{27,28}. However, gonads from female moles exhibited expression of migration markers, such as PDGFR α or MT1-MMP, suggesting that mesonephros-to-gonad migration might contribute to ovotestis formation^{29,30}. However, no signs of migration were observed in the *Sall1*-overexpressing ovaries compared to wildtype ones (**Supplementary Figure 17**), confirming that the expression of additional factors is required to induce cell migration from the mesonephros. Yet, the expression of *SALL1* alone is sufficient to induce kidney-related gene programs, including mesonephros development, which are also observed in ovotestis development. Overall, our findings suggest that the expression of *SALL1* has been co-opted in mole ovotestes formation through the gain of specific enhancers, resulting in the recruitment of tissue-specific transcriptional programs.

DISCUSSION

Across vertebrates, gonadal development is characterized by a remarkable evolutionary plasticity^{31,32}. This is particularly highlighted by the development of ovotestes in moles, in which the development of a testicular region that increases the production of male hormones is fully compatible with a reproductive function³³. In previous studies, we demonstrated that mole ovotestis development is associated with a prolonged expression of *FGF9* through early gonadal development¹¹. This heterochronic expression pattern delays the onset of female meiosis and creates a pro-testicular environment that is critical for ovotestis development. Our transgenic experiment revealed that *SALL1* overexpression contributes to this transcriptional environment by activating gene expression programs. These programs are characterized by molecular

signatures that are shared with other *SALL1*-expressing tissues, such as the kidney. Yet, this ectopic program is not sufficient to trigger sex-reversal mechanisms, as denoted in phenotypical analyses. Therefore, it is plausible that *SALL1* may cooperate with other factors in ovotestis development and/or benefit from the pro-testicular environment that *FGF9* misexpression induces.

During evolution, genes are frequently co-opted for species-specific processes. These effects are often mediated by changes in the activity of regulatory elements that preserve the essential function of genes and, at the same time, allow a diversification of its expression in new tissues and cell types⁶⁻¹⁰. Our analyses showed that mole *SALL1* enhancers were not able to recapitulate gonadal expression in mouse reporter assays. This could indicate that additional trans-acting factors are required for their activation, such as *IRF4* or *FOXP1*. Our analyses revealed that the mole enhancer sequences contain specific binding sites for these transcription factors, which also have higher levels of expression in ovotestes compared to mouse gonads. However, *SALL1* enhancers also display consistent activity in the mesonephros, a tissue that shares a common molecular origin with the gonad. Furthermore, the mesonephros is a known source of endothelial, myoid and supporting cells to the gonad^{34,35}. Interestingly, the developing ovotestes of the mole, in contrast to female gonads of most mammalian species, show a prominent expression of migration markers^{29,30}. Thus, the formation of mole ovotestis may involve the recruitment of cells from the adjacent mesonephros, which may explain the activity of mole enhancers in this tissue. Interestingly, the mesonephric activation of *SALL1* is driven by several enhancers, thus resembling a functional redundancy of CREs that has been described at multiple developmental loci³⁶. Such cooperative activity has been proposed to arise by an initial gain of transcription factor binding sites that is progressively stabilized through the recruitment of additional sites at other elements, giving capacity to these elements to evolve redundant functions³⁷. We observed a similar mechanism in the regulatory landscape of *SALL1*, where several of the enhancers share binding sites for transcription factors specifically expressed in the testicular part of the ovotestes. Furthermore, this pattern of transcription factor binding is highly distinct from other mammal species, such as mouse or shrew, suggesting the capacity of these non-coding elements to evolve.

TAD structures serve as a spatial scaffold, in which regulatory elements interact with their cognate genes, thus representing the existence of large 3D regulatory landscapes contributing to the specificity of gene expression. These domains have been suggested to represent a fertile ground for the evolution of gene expression³⁸⁻⁴⁰. Previous studies have demonstrated that TADs impose important constraints during evolution, as genomic

rearrangements are more prone to occur at boundaries, preserving TADs as entire regulatory units⁴¹. However, genomic rearrangements that reorganize TADs can be also associated with changes in gene expression that might induce the evolution of traits⁴². This has been recently exemplified with the ectopic activation of the PCP pathway linked to the development of enlarged fins in skates, but also in moles where genomic rearrangements affecting the *FGF9* and *CYP17A1* TADs are associated with intersexuality^{11,43}. In contrast, our current study also highlights that the evolution of CREs within conserved TADs is another relevant mechanism for evolution. This is denoted by the striking conservation of TAD organization at the *SALL1* TAD, which is characterized by a remarkable internal evolution of CREs. These results are consistent with previous observations and further reinforce the idea that TADs might serve as a scaffold for the evolution of gene pleiotropy^{44,45}. In summary, our results suggest the co-option of *SALL1* in mole ovotestis development, through regulatory changes that occur despite a striking conservation of TAD organization. This highlights the multilayered nature of gene regulation and how changes at different levels may serve as a driving force for the evolution of traits.

Acknowledgements

We thank Prof. Jiménez lab for hosting and assisting in the mole captures. We thank Asita Stiege, Norbert Brieske, Ute Fisher, Karol Macura, Judith Fiedler, Katja Zill and Chistine Franke for their technical support. We thank Professor Koopman for kindly giving us the *Wt1*-BAC piggybac vector. We thank Prof. Mundlos and Dr. Lupiáñez for their fruitful discussions, as well as all the Mundlos group members for their productive interactions.

Funding

F.M.R. was supported by grants from the Deutsche Forschungsgemeinschaft (MU 880/16-1). M.O. was supported by Swiss National Science Foundation (SNSF) grant PCEFP3_186993. A.V. was supported by National Institutes of Health (NIH) grants R01HG003988, UM1HG009421, R01HL162304, and R01DE028599. Research was conducted at the E.O. Lawrence Berkeley National Laboratory and performed under Department of Energy Contract DE-AC02-05CH11231, University of California.

Competing interests

The authors declare no conflict of interest.

Data availability

Raw and processed datasets used in this study are available in the Gene Expression Omnibus repository under the accession number GSE219025.

References

1. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* **167**, 1170–1187 (2016).
2. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
3. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
4. Wray, G. A. The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.* **8**, 206–216 (2007).
5. Lopez-Rios, J. *et al.* Attenuated sensing of SHH by Ptch1 underlies evolution of bovine limbs. *Nature* **511**, 46–51 (2014).
6. Sanetra, M., Begemann, G., Becker, M.-B. & Meyer, A. Conservation and co-option in developmental programmes: the importance of homology relationships. *Front. Zool.* **2**, 15 (2005).
7. McLennan, D. A. The Concept of Co-option: Why Evolution Often Looks Miraculous. *Evol. Educ. Outreach* **1**, 247–258 (2008).
8. Holland, L. Z. Evolution of new characters after whole genome duplications: Insights from amphioxus. *Semin. Cell Dev. Biol.* **24**, 101–109 (2013).
9. Jandzik, D. *et al.* Evolution of the new vertebrate head by co-option of an ancient chordate skeletal tissue. *Nature* **518**, 534–537 (2015).
10. Chuong, E. B., Elde, N. C. & Feschotte, C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **351**, 1083–1087 (2016).
11. Real, F. M. *et al.* The mole genome reveals regulatory rearrangements associated with adaptive intersexuality. *Science* **370**, 208–214 (2020).
12. McLaren, A. Primordial germ cells in the mouse. *Dev. Biol.* **262**, 1–15 (2003).
13. Ramisch, A. *et al.* CRUP: a comprehensive framework to predict condition-specific regulatory units. *Genome Biol.* **20**, 227 (2019).
14. Adrian, T. E. *et al.* Neuropeptide Y distribution in human brain. *Nature* **306**, 584–586 (1983).
15. Körner, M., Waser, B., Thalmann, G. N. & Reubii, J. C. High expression of NPY receptors in the human testis. *Mol. Cell. Endocrinol.* **337**, 62–70 (2011).
16. Sweetman, D. & Münsterberg, A. The vertebrate spalt genes in development and disease. *Dev. Biol.* **293**, 285–293 (2006).
17. Nishinakamura, R. & Takasato, M. Essential roles of Sall1 in kidney development. *Kidney Int.* **68**, 1948–1950 (2005).
18. Kohlhase, J., Wischermann, A., Reichenbach, H., Froster, U. & Engel, W. Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. *Nat. Genet.* **18**, 81–83 (1998).
19. Ma, Y. *et al.* SALL1 expression in the human pituitary-adrenal/gonadal axis. *J. Endocrinol.* **173**, 437–448 (2002).
20. Nicol, B. *et al.* Genome-wide identification of FOXL2 binding and characterization of FOXL2 feminizing action in the fetal gonads. *Hum. Mol. Genet.* **27**, 4273–4287 (2018).

21. Douady, C. J. *et al.* Molecular phylogenetic evidence confirming the Eulipotyphla concept and in support of hedgehogs as the sister group to shrews. *Mol. Phylogenet. Evol.* **25**, 200–209 (2002).
22. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).
23. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser--a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35**, D88–92 (2007).
24. Tilmann, C. & Capel, B. Mesonephric cell migration induces testis cord formation and Sertoli cell differentiation in the mammalian gonad. *Development* **126**, 2883–2890 (1999).
25. Nishinakamura, R. *et al.* Murine homolog of SALL1 is essential for ureteric bud invasion in kidney development. *Development* **128**, 3105–3115 (2001).
26. Zhao, L., Ng, E. T. & Koopman, P. A piggyBac transposon- and gateway-enhanced system for efficient BAC transgenesis. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **243**, 1086–1094 (2014).
27. Martineau, J., Nordqvist, K., Tilmann, C., Lovell-Badge, R. & Capel, B. Male-specific cell migration into the developing gonad. *Curr. Biol. CB* **7**, 958–968 (1997).
28. Capel, B., Albrecht, K. H., Washburn, L. L. & Eicher, E. M. Migration of mesonephric cells into the mammalian gonad depends on Sry. *Mech. Dev.* **84**, 127–131 (1999).
29. Carmona, F. D. *et al.* The spatio-temporal pattern of testis organogenesis in mammals - insights from the mole. *Int. J. Dev. Biol.* **53**, 1035–1044 (2009).
30. Lupiáñez, D. G. *et al.* Pattern and density of vascularization in mammalian testes, ovaries, and ovotestes. *J. Exp. Zool. B Mol. Dev. Evol.* **318**, 170–181 (2012).
31. Jiménez, R. Ovarian organogenesis in mammals: mice cannot tell us everything. *Sex. Dev. Genet. Mol. Biol. Evol. Endocrinol. Embryol. Pathol. Sex Determ. Differ.* **3**, 291–301 (2009).
32. Capel, B. Vertebrate sex determination: evolutionary plasticity of a fundamental switch. *Nat. Rev. Genet.* **18**, 675–689 (2017).
33. Jiménez, R. *et al.* Fertile females of the mole *Talpa occidentalis* are phenotypic intersexes with ovotestes. *Dev. Camb. Engl.* **118**, 1303–1311 (1993).
34. Burgoyne, P. & Palmer, S. Cellular basis of sex determination and sex reversal in mammals. *Gonadal Dev. Funct.* 17–29 (1993).
35. Brennan, J. & Capel, B. One tissue, two fates: molecular genetic events that underlie testis versus ovary development. *Nat. Rev. Genet.* **5**, 509–521 (2004).
36. Osterwalder, M. *et al.* Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243 (2018).
37. Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).
38. Hoencamp, C. *et al.* 3D genomics across the tree of life reveals condensin II as a determinant of architecture type. *Science* (2021) doi:10.1126/science.abe2218.
39. Anania, C. & Lupiáñez, D. G. Order and disorder: abnormal 3D chromatin organization in human disease. *Brief. Funct. Genomics* **19**, 128–138 (2020).
40. Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018).
41. Krefting, J., Andrade-Navarro, M. A. & Ibn-Salem, J. Evolutionary stability of topologically associating domains is associated with conserved gene regulation. *BMC Biol.* **16**, 87 (2018).

42. Ghavi-Helm, Y. *et al.* Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* 1 (2019) doi:10.1038/s41588-019-0462-3.
43. Marlétaz, F. *et al.* The little skate genome and the evolutionary emergence of wing-like fins. *Nature* **616**, 495–503 (2023).
44. Franke, M. & Gómez-Skarmeta, J. L. An evolutionary perspective of regulatory landscape dynamics in development and disease. *Curr. Opin. Cell Biol.* **55**, 24–29 (2018).
45. Acemel, R. D., Maeso, I. & Gómez-Skarmeta, J. L. Topologically associated domains: a successful scaffold for the evolution of gene regulation in animals. *Wiley Interdiscip. Rev. Dev. Biol.* **6**, (2017).
46. McKay, S. J., Vergara, I. A. & Stajich, J. E. Using the Generic Synteny Browser (GBrowse_syn). *Curr. Protoc. Bioinforma.* **Chapter 9**, Unit 9.12 (2010).
47. Barrionuevo, F. J., Zurita, F., Burgos, M. & Jiménez, R. Testis-like development of gonads in female moles. New insights on mammalian gonad organogenesis. *Dev. Biol.* **268**, 39–52 (2004).
48. Lázaro, J., Hertel, M., Sherwood, C. C., Muturi, M. & Dechmann, D. K. N. Profound seasonal changes in brain size and architecture in the common shrew. *Brain Struct. Funct.* **223**, 2823–2840 (2018).
49. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).
50. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
51. Thomas-Chollier, M. *et al.* Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nat. Protoc.* **6**, 1860–1869 (2011).
52. Durand, N. C. *et al.* Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst.* **3**, 99–101 (2016).
53. van de Werken, H. J. G. *et al.* 4C technology: protocols and data analysis. *Methods Enzymol.* **513**, 89–112 (2012).
54. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinforma. Oxf. Engl.* **26**, 589–595 (2010).
55. Durand, N. C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst.* **3**, 95–98 (2016).
56. Osterwalder, M. *et al.* Characterization of Mammalian In Vivo Enhancers Using Mouse Transgenesis and CRISPR Genome Editing. *Methods Mol. Biol. Clifton NJ* **2403**, 147–186 (2022).
57. Wang, J. *et al.* An improved recombineering approach by adding RecA to lambda Red recombination. *Mol. Biotechnol.* **32**, 43–53 (2006).
58. Rostovskaya, M. *et al.* Transposon-mediated BAC transgenesis in human ES cells. *Nucleic Acids Res.* **40**, e150 (2012).
59. Thomas, P. D. *et al.* PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* **13**, 2129–2141 (2003).

Figures

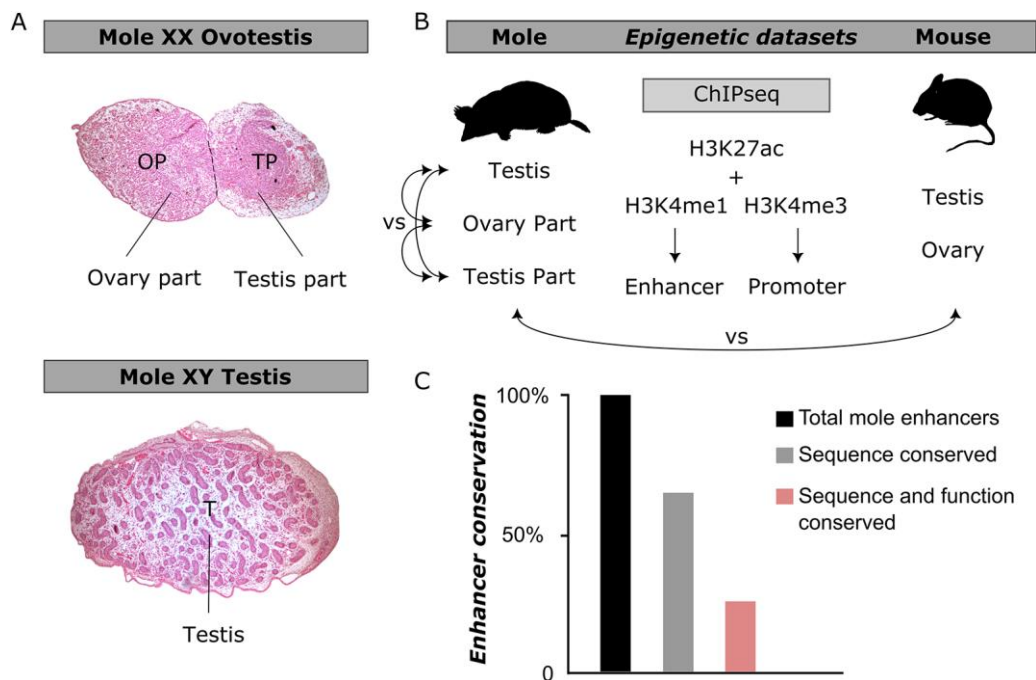


Fig. 1. Characterization of regulatory elements in mole ovotestes.

- Hematoxylin and eosin staining of mole gonads at postnatal stage P7. Female ovotestis in upper panel, testis in lower panel. OP: ovary part, TP: testis part, T: testis. Note the clear separation into two parts of the ovotestis. Scale bars: 100 μ m.
- Scheme of the gonadal tissues sampled to generate the epigenetic datasets in mole and mouse. Five tissues and three different histone modifications were used for the ChIP-seq experiments.
- Percentage of mole enhancers conserved compared to mice. Conservation at the sequence level is shown in gray, conservation at the enhancer signature level is shown in light red.

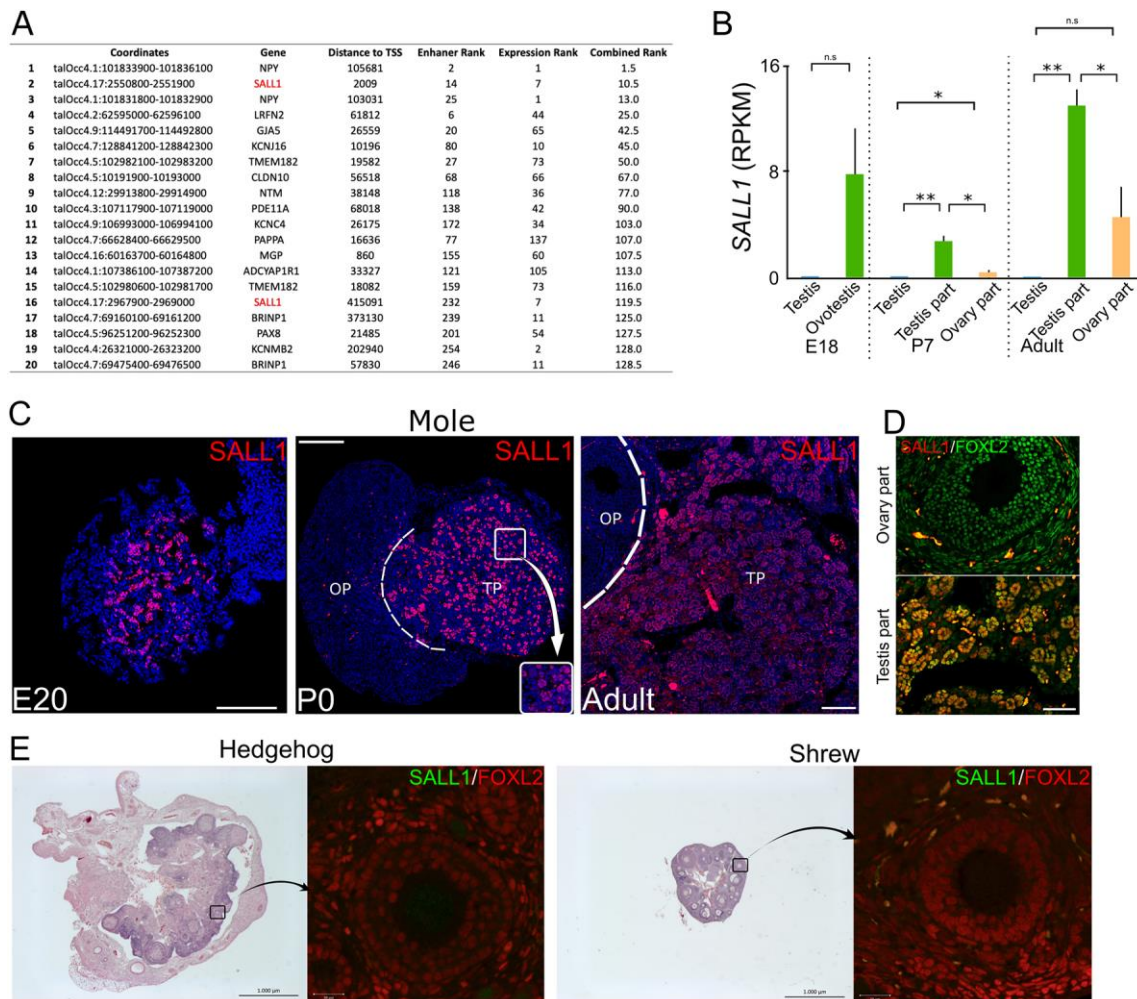


Fig. 2. Identification of *SALL1* as a marker for testis part formation in mole ovotestes.

- Top 20 enhancer regions ranked by enhancer score and specificity of expression of the associated gene in the testis part of the ovotestis. Note two *SALL1* enhancers which are highly ranked (rank 2+16).
- SALL1* expression levels in RPKM (reads per kilobase million) from mole RNA-seq data at different developmental time points.
- Spatio-temporal profile of *SALL1* expression in mole ovotestes (immunofluorescence, *SALL1* in red; DAPI in blue). Note that *SALL1* is spatially restricted to the medullary (testicular) region of the mole ovotestis at E20 and is also present in the testis part thereafter. Inset shows localization to Sertoli-like cells. OP: ovarian part, TP: testicular part. Scale bars: 100 μ m.
- Double immunostaining for *SALL1* and *FOXL2* in adult ovotestes. Note, the absence of *SALL1* positive cells in the ovary part, contrary to the testis part

where both markers co-localize in the spherules (equivalent to testis cords).
Scale bars: 100 μm .

- E. Spatial expression of SALL1 in adult female hedgehog (*Atelerix albiventris*, left) and adult female shrew (*Sorex araneus*, right) (immunofluorescence, SALL1 in green, ovarian marker FOXL2 in red). Note absence of SALL1 expression. Black and white scale bars represent 1000 μm and 20 μm , respectively.

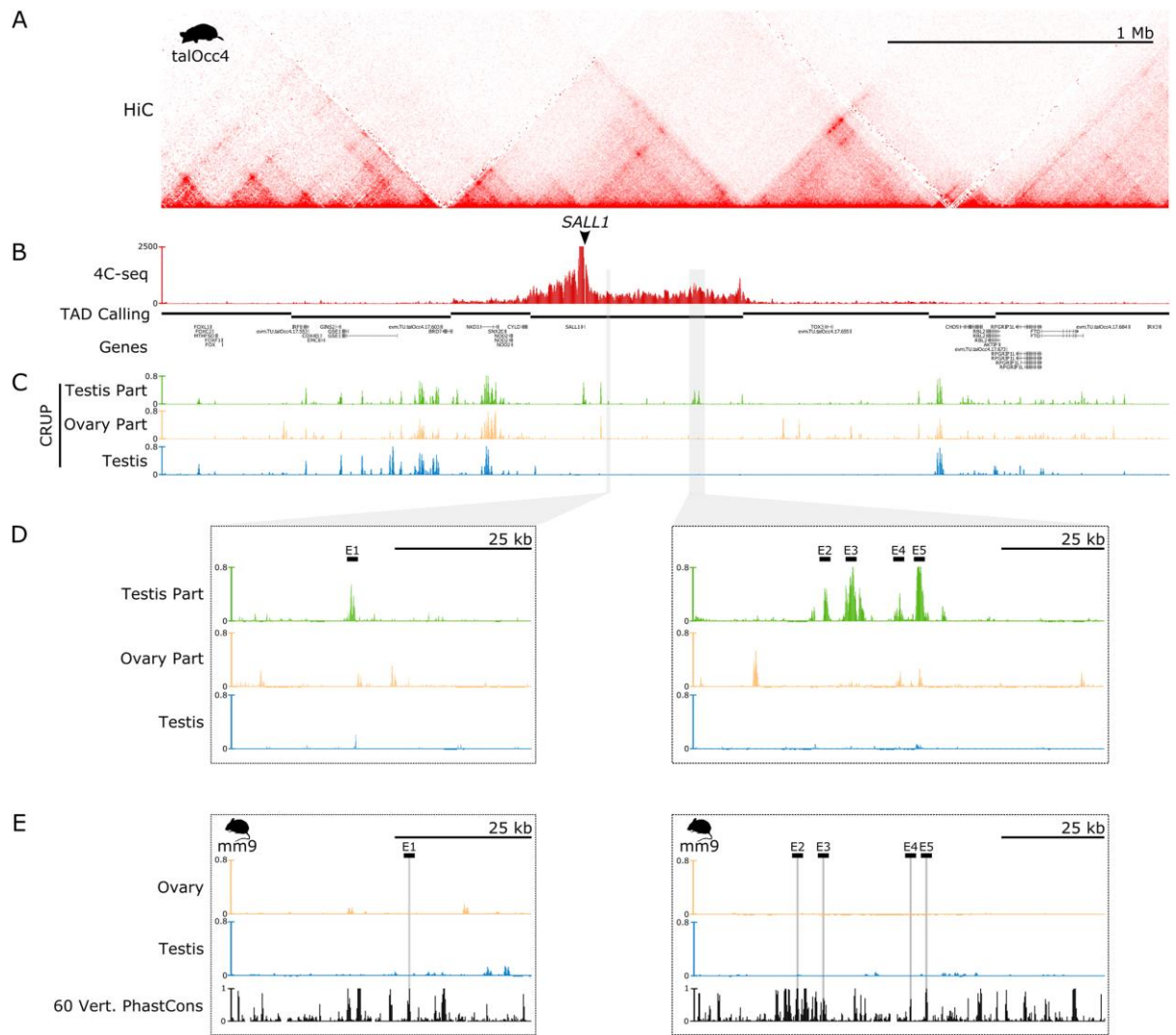


Fig. 3. Regulatory domain and the epigenetic landscape of *SALL1*

- Hi-C map from mole embryonic limbs denotes the domain of *SALL1* in a large gene desert.
- 4C-seq analysis from female adult testis part with *SALL1* promoter as viewpoint. Note high interaction frequency between the gene promoter and the surrounding 1Mb desert clearly demarcating the *SALL1* regulatory domain.
- Epigenetic landscape of *SALL1* in the three tissues sampled, called with the tool CRUP. Note the presence of numerous active enhancers in the testicular part of the ovotestis where *SALL1* is specifically expressed.

- D. Zoom-in on two mole regions containing five specific regulatory elements for the testis part of the ovotestes, named as enhancer 1 to 5 (E1-5).
- E. Homologous regions to the testis part enhancers in the mouse genome. Homologous regions are marked as gray bars. Note the absence of enhancer activity of these regions.

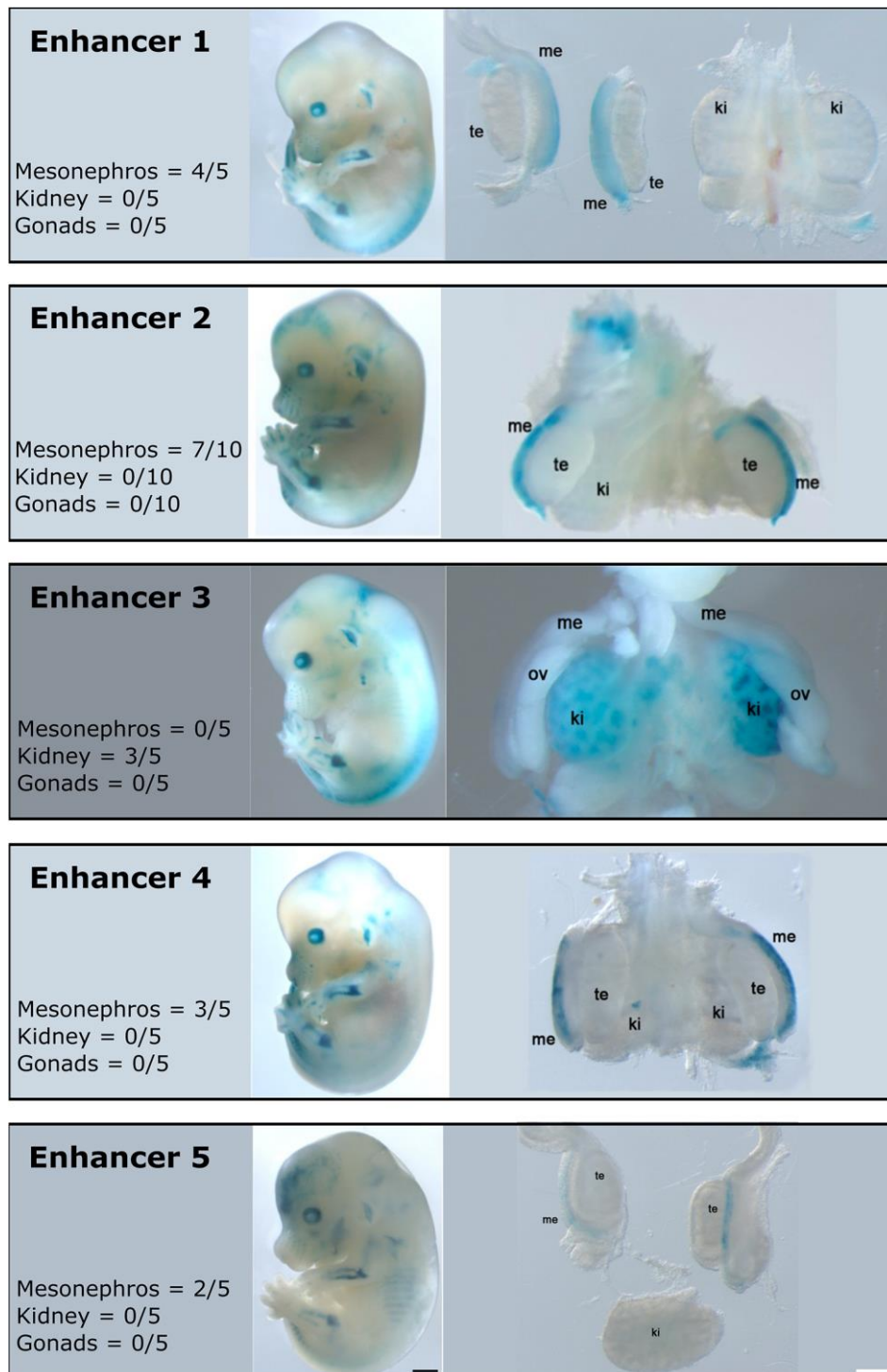


Fig. 4. LacZ reporter assays for enhancer elements E1-5 associated with *SALL1*

The enhancer activity of each element is depicted in separated boxes 1 to 5. Entire embryos at E13.5 as well as the dissected urogenital tracts are displayed. Me: mesonephros, te: testes, ov: ovaries, ki: kidneys. Black and white scale bars represent 1000 and 100 μm , respectively.

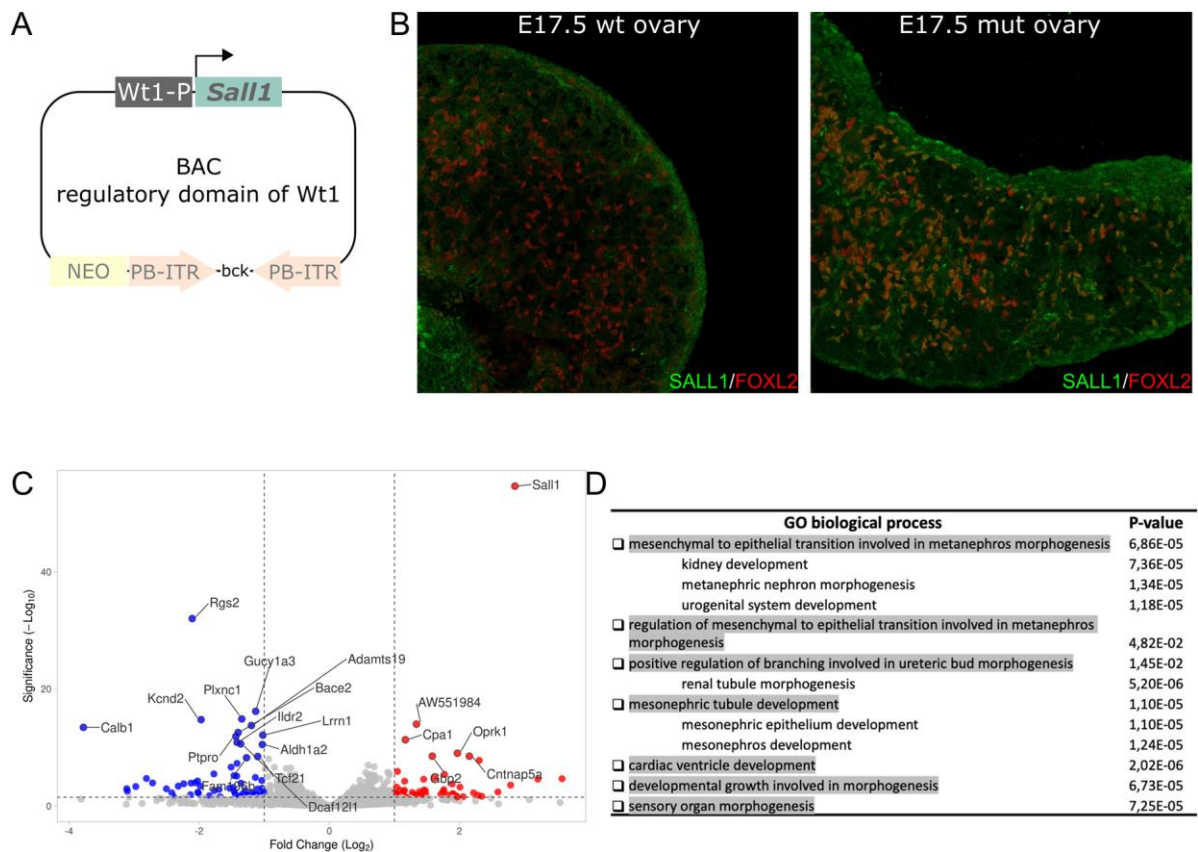


Fig. 5. Overexpression of *Sall1* in mouse embryonic ovaries results in hundreds of deregulated genes.

- A.** Cloning strategy to overexpress *Sall1* in somatic ovarian cells through BAC transgenesis. *Sall1* is regulated under the promoter and regulatory regions of the gonadal somatic gene, *Wt1*.
- B.** Immunostaining against SALL1 (green) and FOXL2 (red) in wildtype and mutant ovaries at E17.5. Note the high abundance of double positive cells (orange) for SALL1 and FOXL2 in the mutant gonad, confirming the overexpression success.
- C.** Volcano plot from RNA-seq of mutant ovaries compared to control ovaries from littermates at E13.5. Names of the 20 most deregulated genes are indicated. Note that *Sall1* is the most significantly upregulated gene, as shown in the upper right corner of the plot. X axis shows the expression changes in Log₂ fold-change and Y axis shows the p-value.
- D.** Gene ontology enrichment analyses of the common upregulated genes in the *Sall1* mutant ovaries and in the testis part of the ovotestes. Shown is the enriched biological process as well as the P-value.

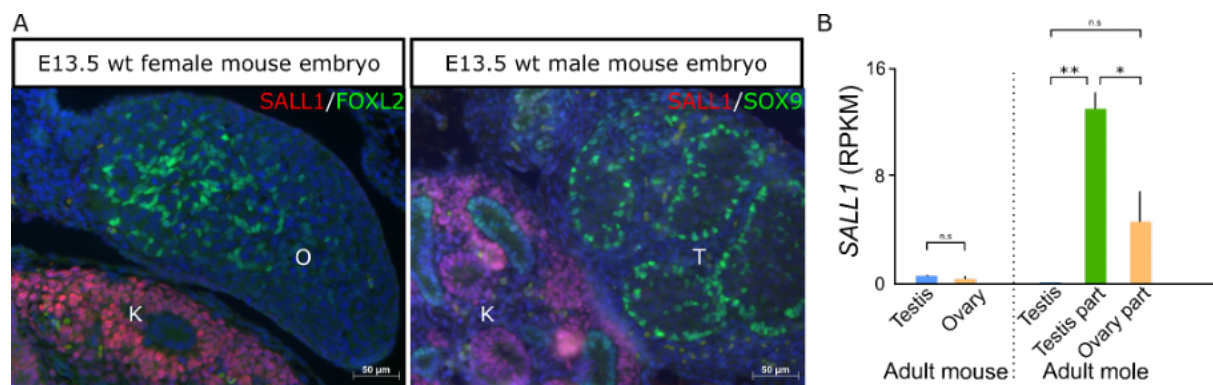


Fig. S1. *Sall1* expression pattern in mouse gonads.

- A. Immunostainings of SALL1 (red) and FOXL2 and SOX9 (green) as markers of somatic female and male cells, respectively. O: ovary, T: testis, K: kidney. Note the absence of SALL1 positive cells in the embryonic gonads but the specific expression in the adjacent kidney. Scale bars: 50 µm.
- B. RPKMs quantification from RNA-seq data of adult gonads in mouse and mole. Expression levels in mouse are lower compared to mole and not sex specific.

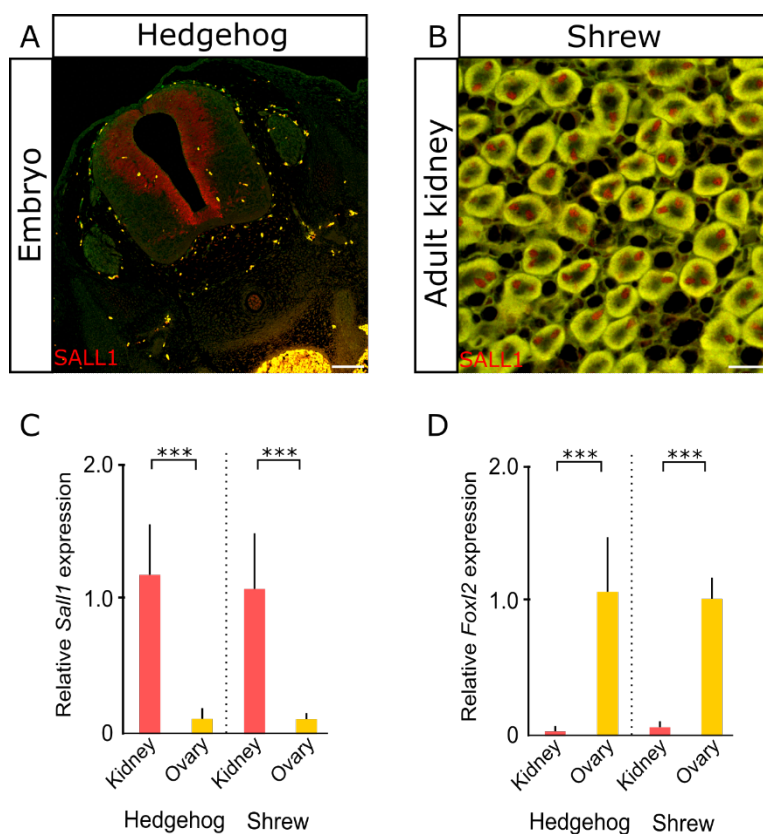


Fig. S2. *SALL1* expression in *Eulipotyphla* species.

A. Immunostaining of SALL1 in transversal sections of an early hedgehog embryo from the *Atelerix albiventris* species. SALL1 is highly expressed in the neural tube, a well-known tissue for SALL1 expression. Scale bar: 100 μ m.

B. Immunostaining of SALL1 in adult kidneys from the common shrew, *Sorex araneus*. Note the specificity of the antibody to the nucleus of the renal tubular cells. Scale bar: 20 μ m.

C. RT-qPCRs for *SALL1* expression in adult ovaries and kidneys from hedgehogs (*Atelerix albiventris*) and shrews (*Sorex araneus*). Shown is relative *SALL1* expression normalized to *RPS9*. Data is presented as mean \pm SD and p-values are indicated as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

D. RT-qPCRs for *FOXL2* expression in adult ovaries and kidneys from hedgehogs (*Atelerix albiventris*) and shrews (*Sorex araneus*). Shown is relative *FOXL2* expression normalized to *RPS9*. Data is presented as mean \pm SD and p-values are indicated as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

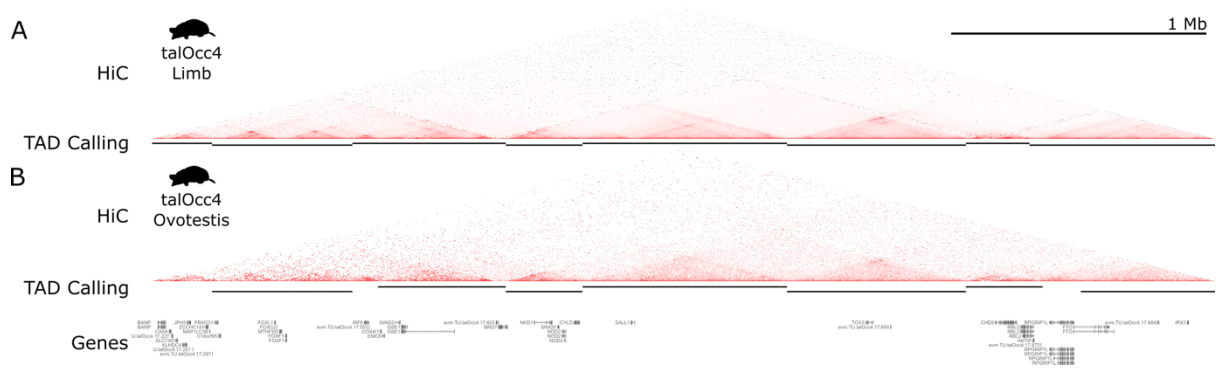


Fig. S3. Hi-C map comparison between limb and ovotestis

- Hi-C maps at high resolution from embryonic limbs with the corresponding TAD calling (black bars) underneath.
- Hi-C maps from adult ovotestis with the corresponding TAD calling (black bars) underneath. Note the conservation of the *SALL1* TAD domain between tissues.

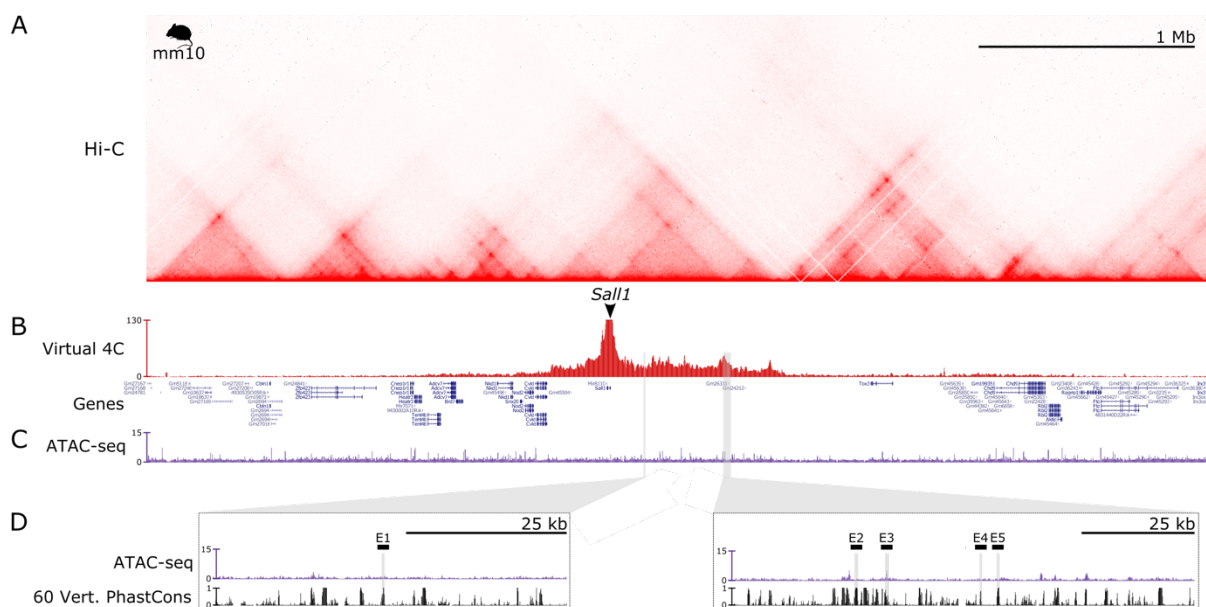
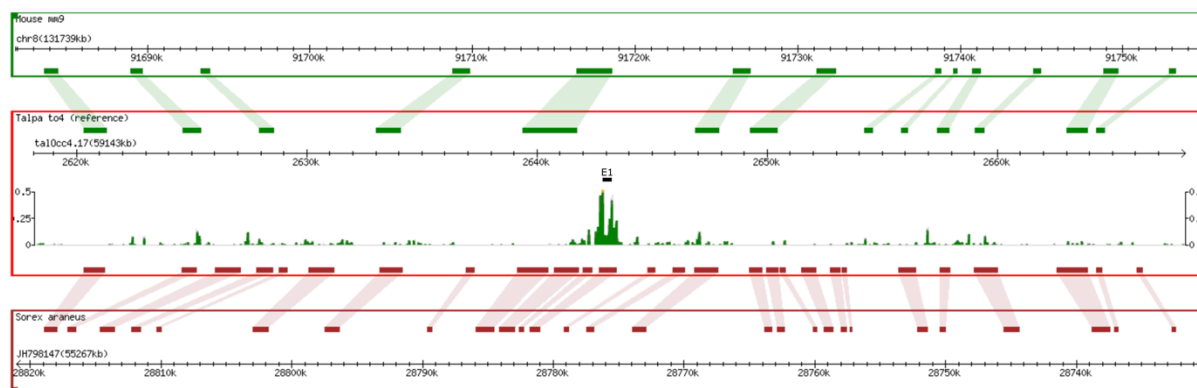


Fig. S4. Regulatory domain of *Sall1* in mouse.

- A. Hi-C map from Neural Progenitor Cells (NPCs) denotes the domain of *Sall1* in a large gene desert.
- B. Virtual 4C-seq analysis from NPCs Hi-C maps with *SALL1* promoter as viewpoint. Note high interaction frequency between the gene promoter and the surrounding 1Mb desert clearly demarcating the *Sall1* regulatory domain. The domain is strikingly conserved between cell types and species.
- C. ATAC-seq track from mouse embryonic kidneys at E14.5 to identify regulatory regions in this tissue.
- D. Zoom-in on the two equivalent regions where the mole enhancers were identified. Homologous regions are marked as gray bars and labeled as E1-5. Consistent with our enhancer activity results, enhancer 3 (E3) coincides with an ATAC-seq peak in kidneys.

A



B

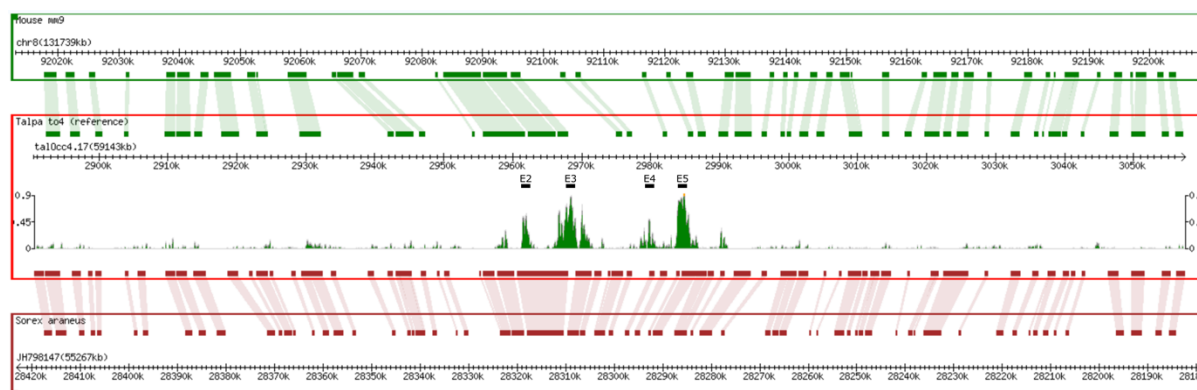


Fig. S5. Synteny of the enhancer regions.

A. Alignment of syntenic blocks for the enhancer region 1 (E1) against the mouse genome (upper panel) and against the shrew genome (*Sorex araneus*, lower panel). Visualization with Gbrowse⁴⁶.

B. Alignment of syntenic blocks for the cluster of enhancers (E2-E5) against the mouse genome (upper panel) and against the shrew genome (*Sorex araneus*, lower panel). Visualization with Gbrowse⁴⁶.

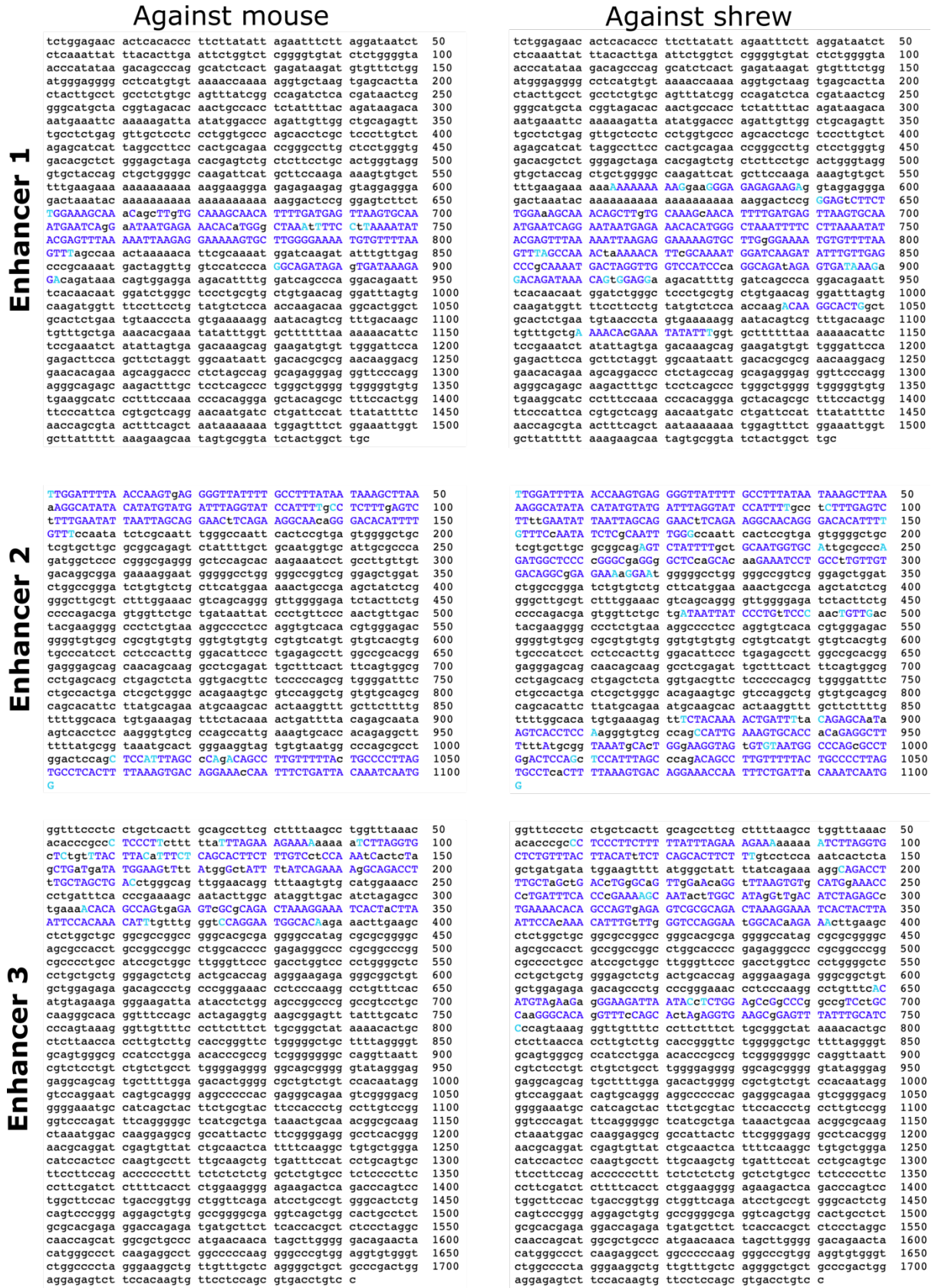


Fig. S6. Sequence alignments for the individual enhancers 1 to 3.

The conserved nucleotides are highlighted in blue and capitalized. Light blues nucleotides denote the beginning and end of the homology sequence.



Fig. S7. Sequence alignments of the individual enhancers 4 and 5.

The conserved nucleotides are highlighted in blue and capitalized. Light blue nucleotides denote the beginning and end of the homology sequence. Note that for enhancer 4 there was no homology in the sequence compared to mouse, human was used instead.

Enhancer 1

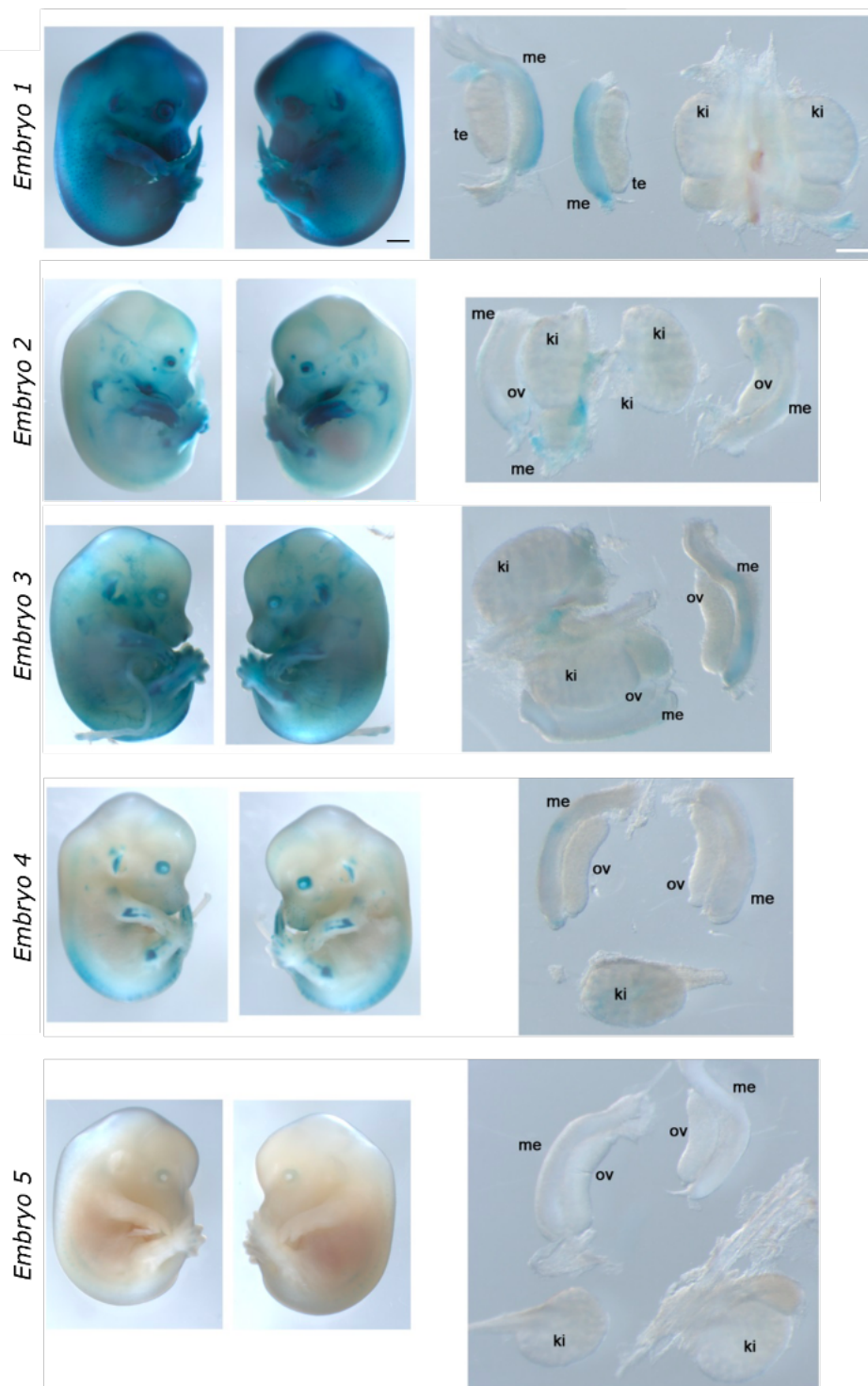


Fig. S8. LacZ enhancer reporter assay for Enhancer 1.

All embryos analyzed for this enhancer are depicted. Entire embryos at E13.5 as well as the dissected urogenital tracts are displayed. me: mesonephros, te: testes, ov: ovaries, ki: kidneys. Four out of five embryos showed mesonephros-specific staining. Black scale bars: 1000 μ m, white scale bars: 100 μ m.

Enhancer 2

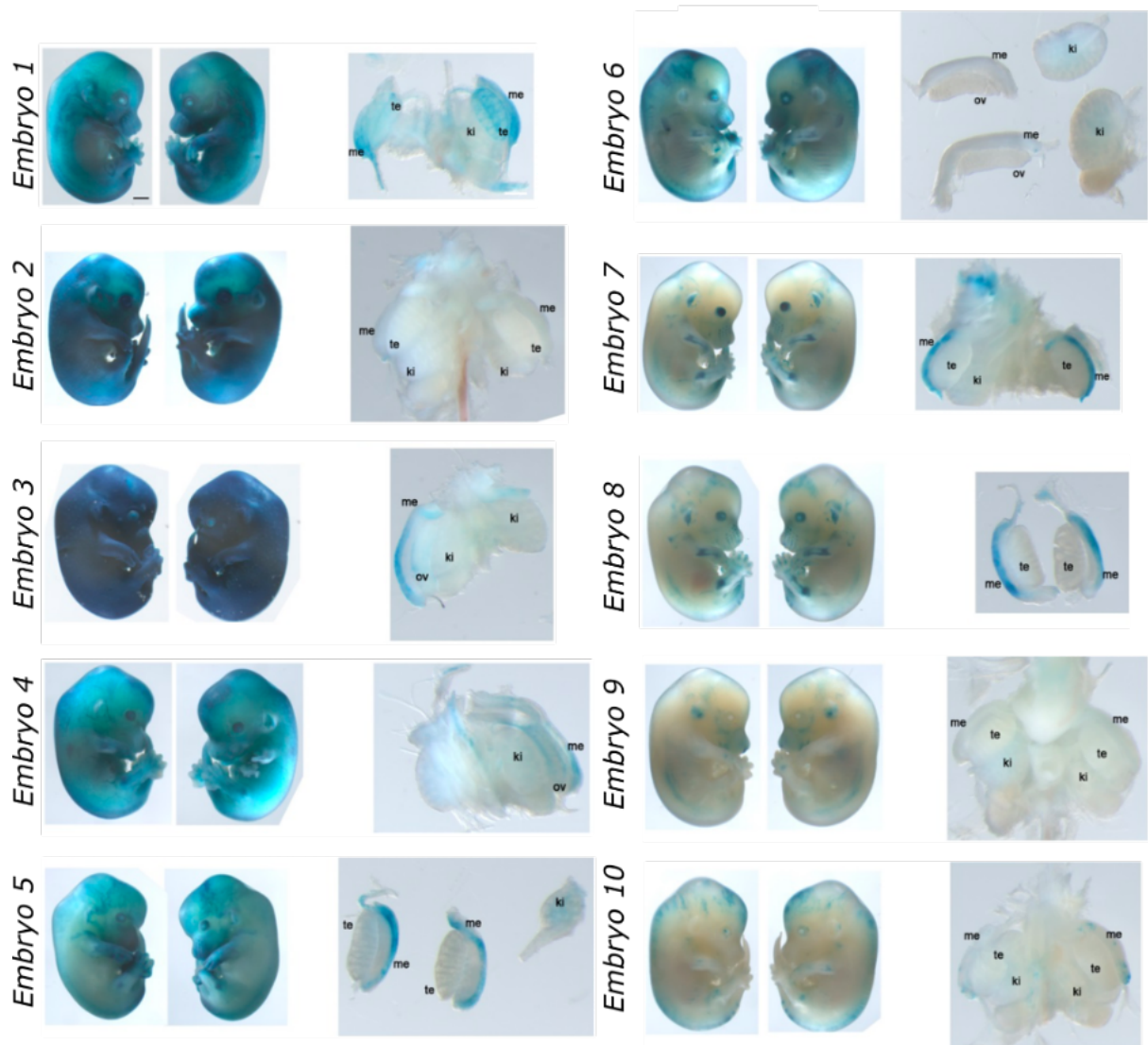


Fig. S9. LacZ enhancer reporter assay for Enhancer 2.

All embryos analyzed for this enhancer are depicted. Entire embryos at E13.5 as well as the dissected urogenital tracts are displayed. Me: mesonephros, te: testes, ov: ovaries, ki: kidneys. Seven out of ten embryos showed mesonephros-specific staining. Black scale bars: 1000 μ m, white scale bars: 100 μ m.

Enhancer 3

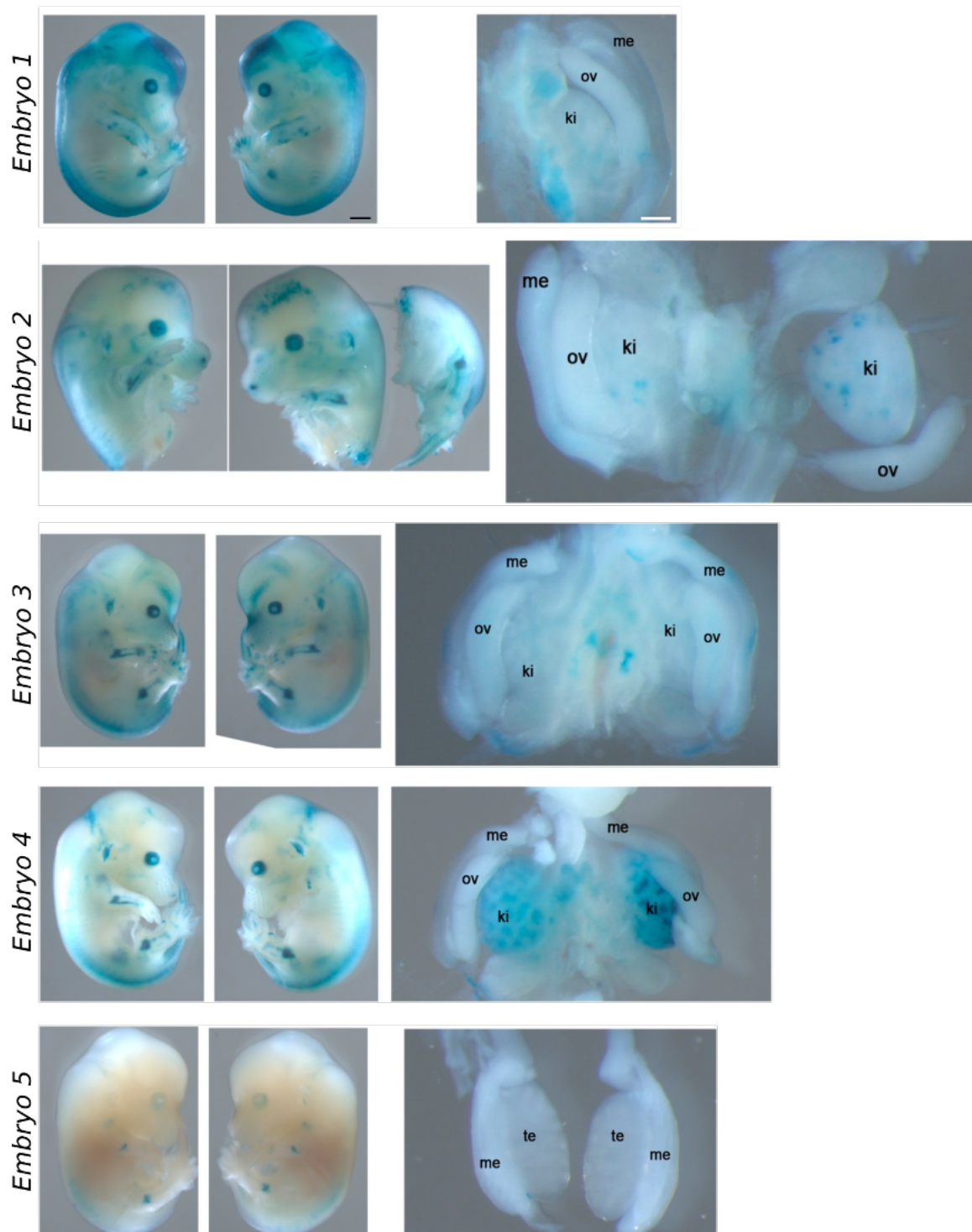


Fig. S10. LacZ enhancer reporter assay for Enhancer 3.

All embryos analyzed for this enhancer are depicted. Entire embryos at E13.5 as well as dissected urogenital tracts are displayed. Me: mesonephros, te: testes, ov: ovaries, ki: kidneys. Three out of five embryos showed kidney-specific staining. Black scale bars: 1000 μm , white scale bars: 100 μm .

Enhancer 4

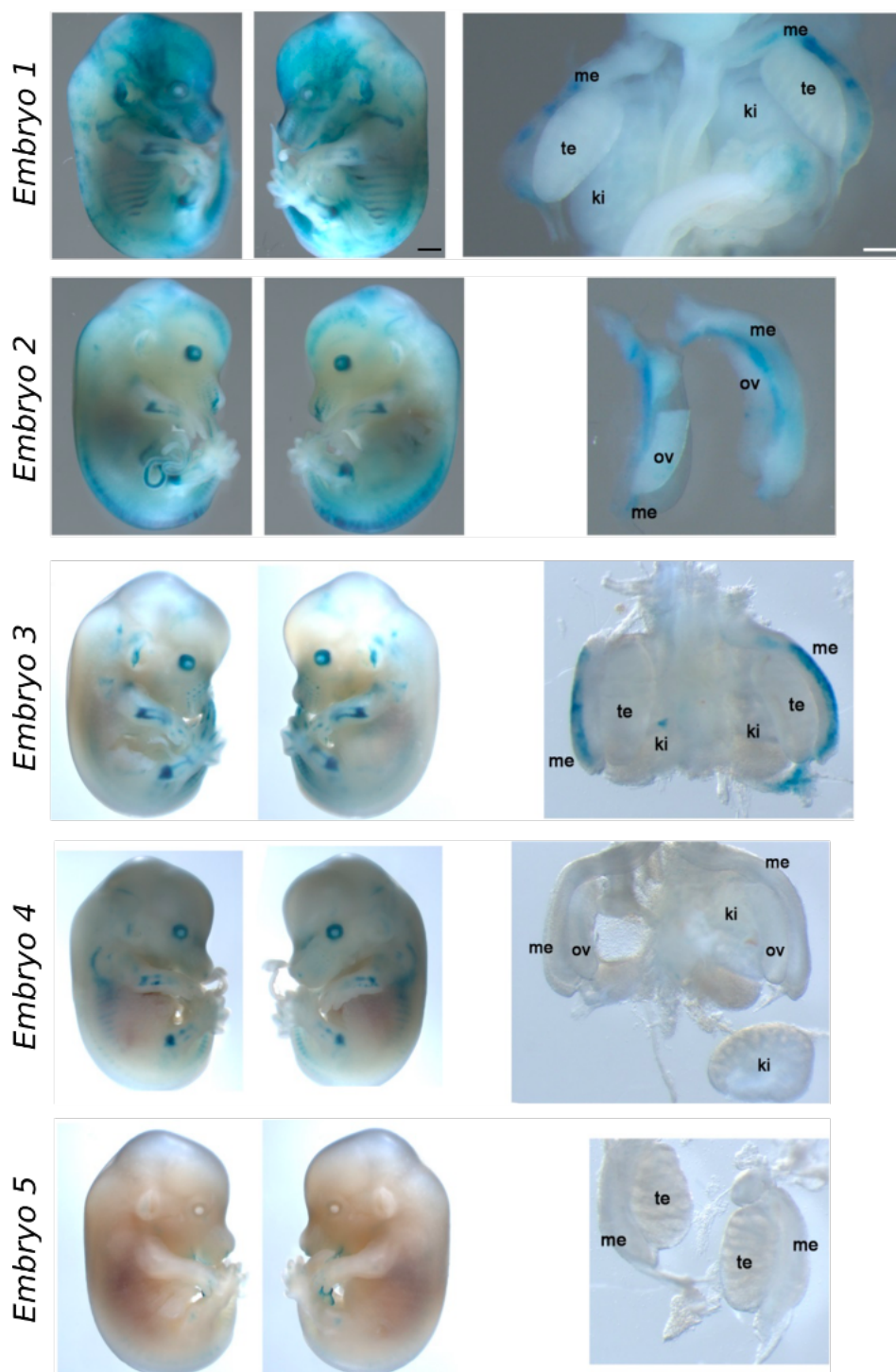


Fig. S11. LacZ enhancer reporter assay for Enhancer 4.

All embryos analyzed for this enhancer are depicted. Entire embryos at E13.5 as well as the dissected urogenital tracts are displayed. Me: mesonephros, te: testes, ov: ovaries, ki: kidneys. Three out of five embryos showed mesonephros-specific staining. Black scale bars: 1000 μ m, white scale bars: 100 μ m.

Enhancer 5

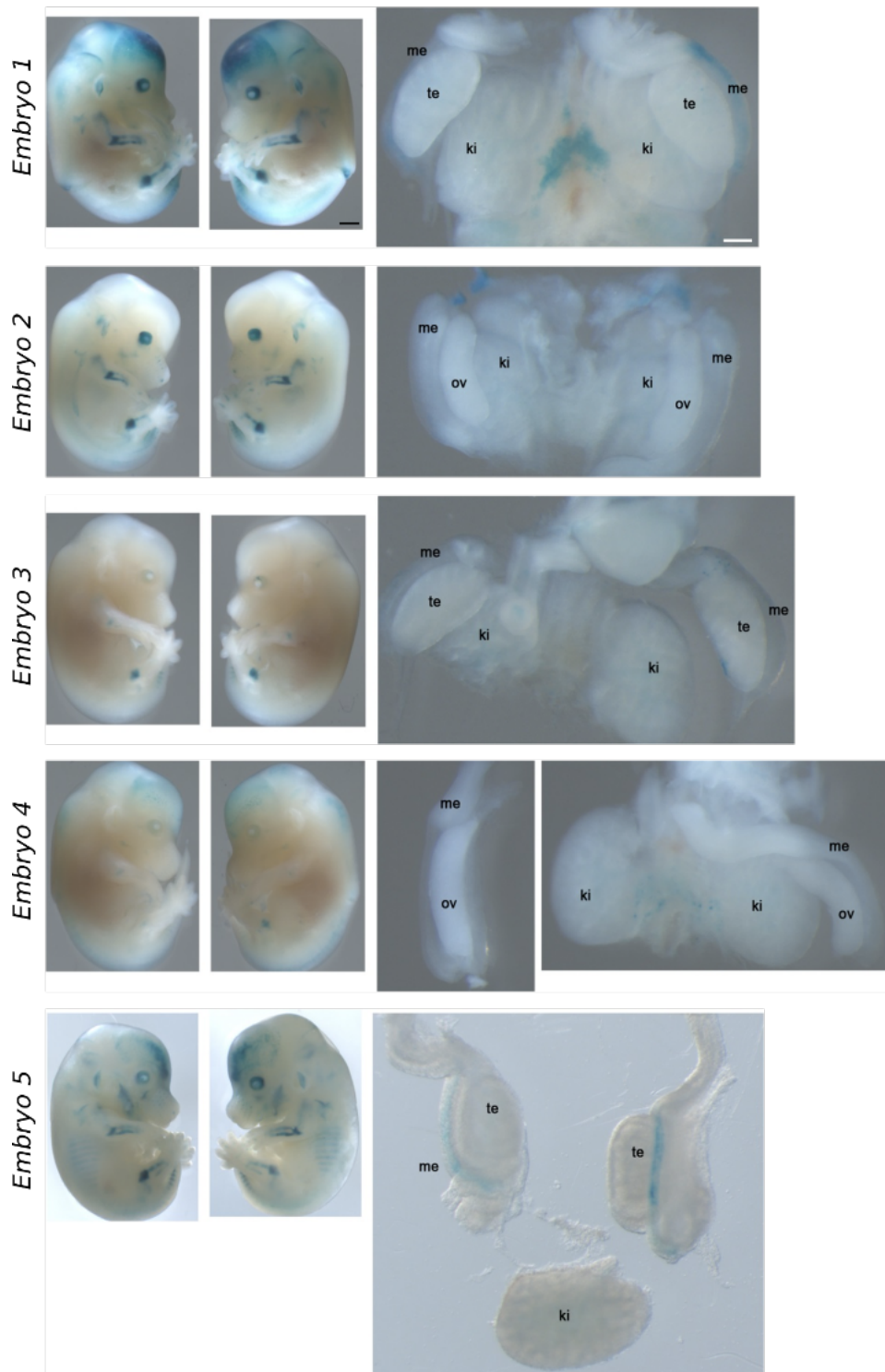


Fig. S12. LacZ enhancer reporter assay for Enhancer 5.

All embryos analyzed for this enhancer are depicted. Entire embryos at E13.5 as well as the dissected urogenital tracts are displayed. Me: mesonephros, te: testes, ov: ovaries, ki: kidneys. Two out of five embryos showed mesonephros-specific staining. Black scale bars: 1000 μm , white scale bars: 100 μm .

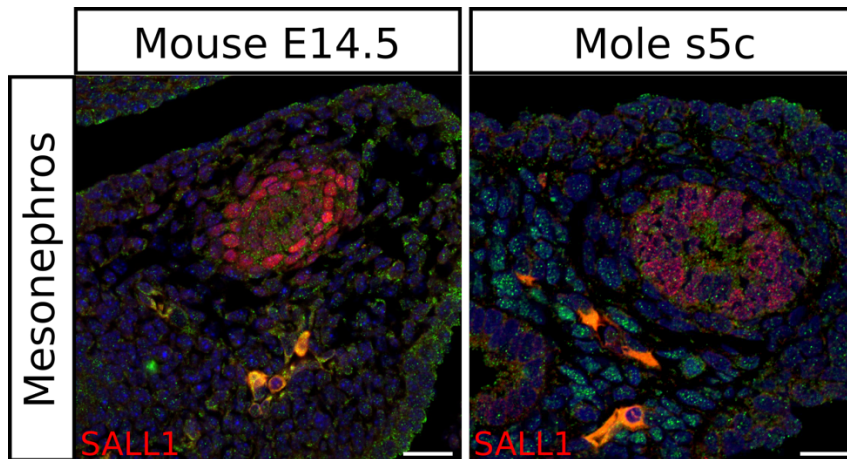


Fig. S13. SALL1 expression in mesonephros

SALL1 is detected in the mesonephros duct of mouse at E14.5 and at equivalent stages in moles (s5c). Scale bars: 20 μ m.

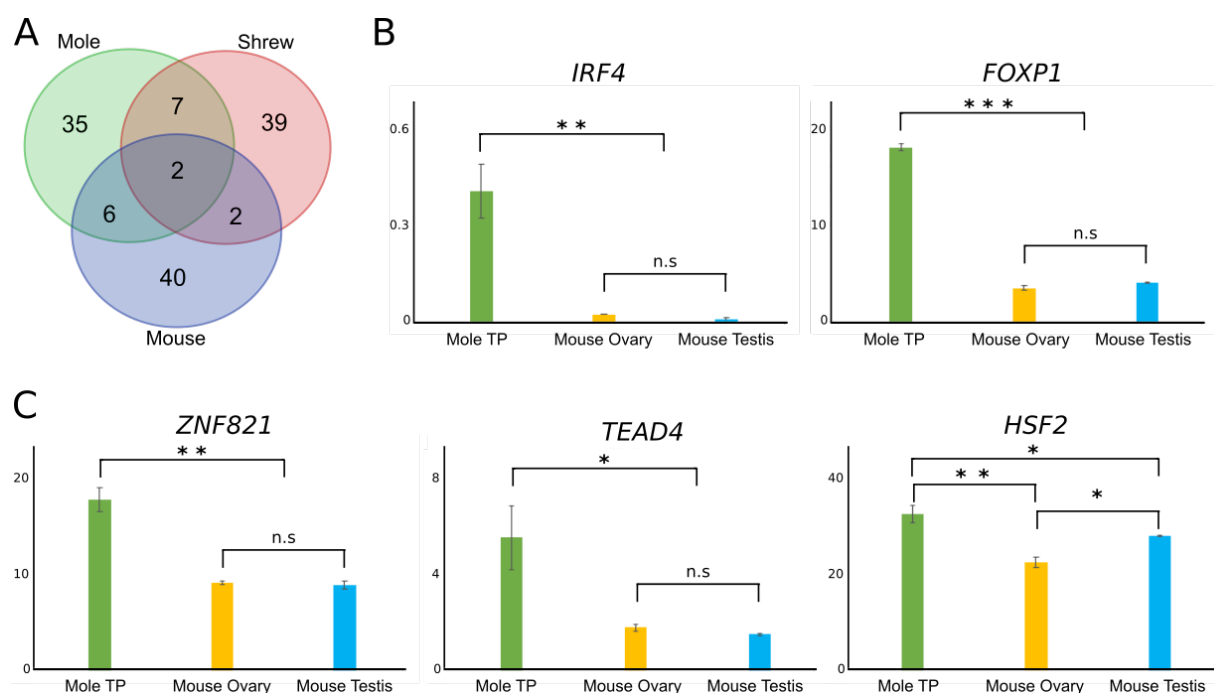


Fig. S14. Comparative analyses of transcription factor binding motifs and expression.

A. Venn diagram showing the number of shared transcription factor binding motifs among the top 50 motifs found in mole, shrew and mouse sequences (**Supplementary Table 2**). Note the limited conservation, emphasizing the sequence divergence observed among species.

B, C. Expression levels in RPKM of transcription factors with top-ranked motif bindings sites in the mole enhancer sequences. The mole TP (testicular part) of the female ovotestis at P7 is compared with the mouse ovary and testis at E13.5. Note the upregulation of these 5 transcription factors when compared to mouse gonads. Data is presented as mean \pm SD and p-values are indicated as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

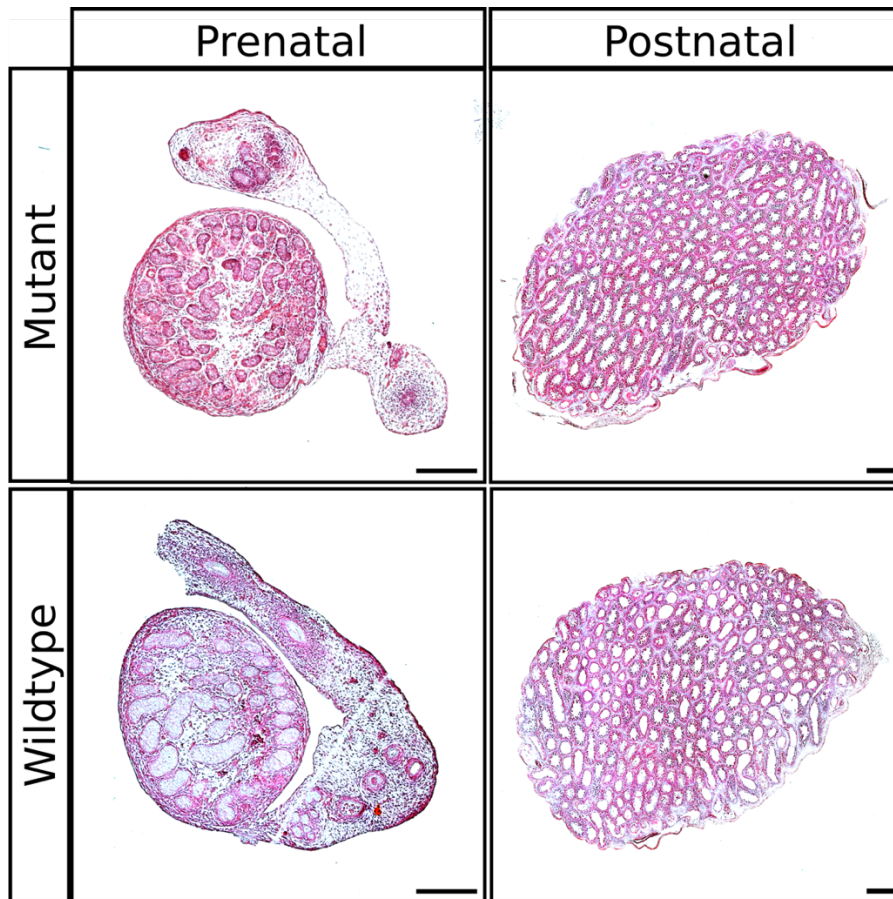


Fig. S15. Morphology of *Sall1*-overexpressing testes during gonad development.

Hematoxylin-eosin staining of mutant overexpressing-*Sall1* and wildtype controls testes before and after birth. There are no differences in size, tissue structure or cell composition between mutants and controls. Scale bars: 200 μ m.

A

	Mus musculus (REF)		upload_1 (▼ Hierarchy NEW! ⓘ)				
GO biological process complete	#	#	expected	Fold Enrichment	+/-	raw P value	FDR
metanephric nephron morphogenesis	18	3	.04	67.89	+	1.79E-05	4.03E-02
↳metanephros morphogenesis	24	3	.06	50.92	+	3.89E-05	4.09E-02
↳animal organ development	3306	20	8.12	2.46	+	6.24E-05	4.68E-02
↳system development	3813	25	9.36	2.67	+	8.56E-07	1.35E-02
↳multicellular organism development	4559	26	11.19	2.32	+	1.01E-05	2.64E-02
↳urogenital system development	368	7	.90	7.75	+	3.46E-05	4.20E-02
↳animal organ morphogenesis	1057	11	2.59	4.24	+	4.59E-05	4.02E-02
↳anatomical structure morphogenesis	2337	19	5.74	3.31	+	1.50E-06	1.18E-02
↳nephron development	136	5	.33	14.98	+	2.45E-05	3.51E-02
positive regulation of branching involved in ureteric bud morphogenesis	23	3	.06	53.13	+	3.47E-05	3.90E-02
↳regulation of branching involved in ureteric bud morphogenesis	25	3	.06	48.88	+	4.35E-05	4.04E-02
↳regulation of morphogenesis of a branching structure	66	4	.16	24.69	+	2.61E-05	3.43E-02
↳regulation of multicellular organismal process	2989	19	7.34	2.59	+	5.28E-05	4.38E-02
ventricular septum development	82	4	.20	19.87	+	5.88E-05	4.63E-02
↳cardiac ventricle development	145	6	.36	16.86	+	1.84E-06	9.66E-03
↳cardiac chamber development	192	6	.47	12.73	+	8.77E-06	2.76E-02
developmental growth involved in morphogenesis	153	5	.38	13.31	+	4.22E-05	4.16E-02
anatomical structure formation involved in morphogenesis	980	11	2.41	4.57	+	2.32E-05	3.65E-02
negative regulation of multicellular organismal process	1154	12	2.83	4.24	+	1.96E-05	3.87E-02
epithelium development	1161	12	2.85	4.21	+	2.08E-05	3.65E-02
↳tissue development	1785	16	4.38	3.65	+	3.88E-06	1.53E-02

B

	Mus musculus (REF)		upload_1 (▼ Hierarchy NEW! ⓘ)				
GO cellular component complete	#	#	expected	Fold Enrichment	+/-	raw P value	FDR
extracellular region	2875	20	7.06	2.83	+	7.91E-06	1.61E-02

Fig. S16. Gene ontology enrichment of commonly upregulated genes in female mole testis part and mouse *Sall1*-overexpressing mutant ovaries.

- GO terms for biological processes.
- GO terms for cellular components.

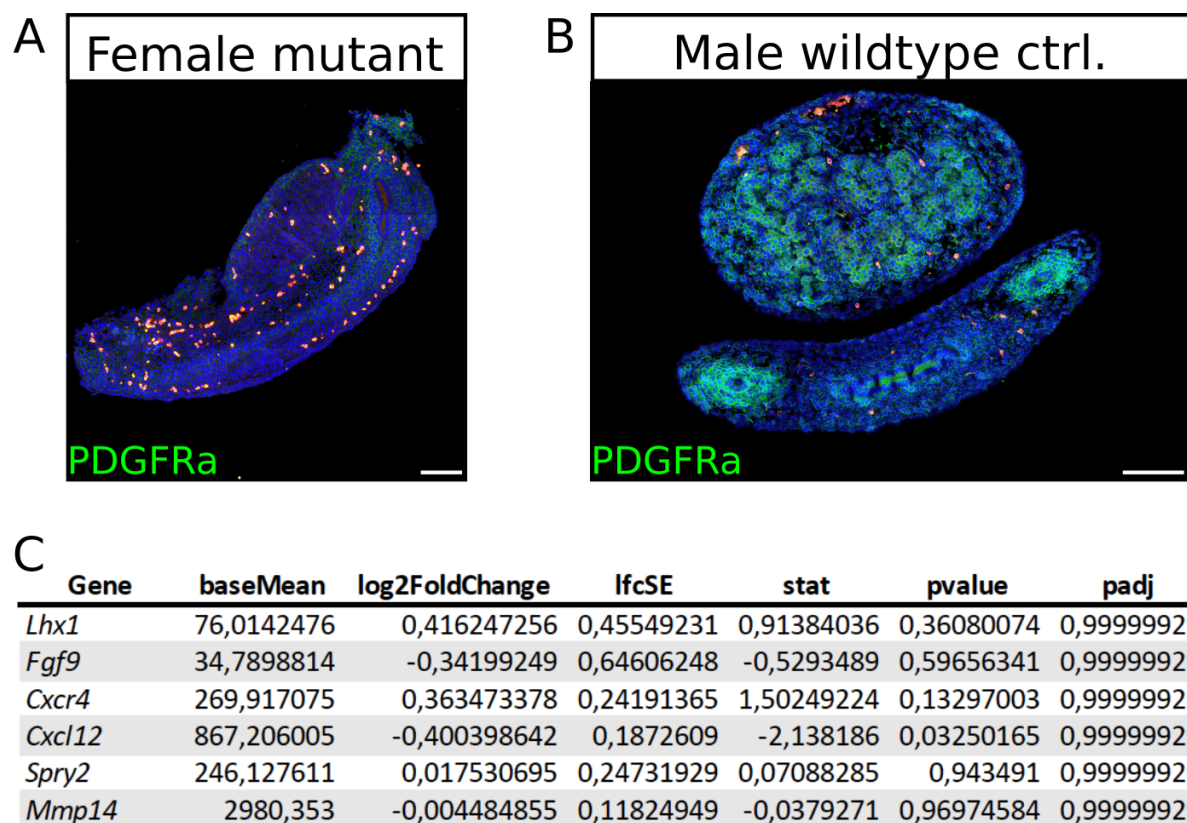


Fig. S17. Expression of migration markers in *Sall1*-overexpressing ovaries.

A, B. Immunostaining for PDGFRa in female mutant and male wildtype controls at E13.5. There is no signal for PDGFRa in mutant ovaries, denoting the absence of migration. Scale bars: 100 μ m

C. Differential gene expression between female mutant and controls for several genes involved in cell migration. Note, there are no significant differences between both conditions.

Table S1. Ranking of enhancer regions. (Provided as an excel table).

[Click here to download Table S1](#)

Table S2. Ranking of transcription factors by significance of binding affinity to the five *SALL1* enhancer sequences. (Provided as an excel table).

[Click here to download Table S2](#)

Table S3. Differential gene expression between *Sall1*-overexpressing mutant and wildtype ovaries. (Provided as an excel table).

[Click here to download Table S3](#)

Table S4. Primer list.

RT-qPCRs	
Hedgehog-qPCR-Sall1-Fwd	GAAGCAAGCGAAGCCTCAAC
Hedgehog-qPCR-Sall1-Rev	TGCTCTTAGTGGGGCGATT
Hedgehog-qPCR-Foxl2-Fwd	CAGAAGCCGCCCTATTCGT
Hedgehog-qPCR-Foxl2-Rev	GGGAACTTGCGGATGATGT
Hedgehog-qPCR-Rps9-Fwd	GCCAAGTCCATCCACCAC
Hedgehog-qPCR-Rps9-Rev	CCAGGCGGACAATGAAGG
Shrew-qPCR-Sall1-Fwd	AGAGCGTTCACAACAAAAGG
Shrew-qPCR-Sall1-Rev	TGGGGCCATCCACAGAGA
Shrew-qPCR-Foxl2-Fwd	CATCGCCAAGTCCCCTTCT
Shrew-qPCR-Foxl2-Rev	GCACTCGTTGAGGCTGAGGT
Shrew-qPCR-Rps9-Fwd	GAGTCCAGGCGAACAATGAA
Shrew-qPCR-Rps9-Rev	GGCCAAGTCCATCCACCA

4C-seq experiments	
Sall1-4C-Fwd	TCAGTGGGCTGACATTTTA
Sall1-4C-Rev	TCAGTGGGCTGACATTTTA
5ITR-4C-Fwd	gctgcacctacagtttggat
5ITR-4C-Rev	gctgcacctacagtttggat
3ITR-4C-Fwd	gctgcacctacagtttggat
3ITR-4C-Rev	gctgcacctacagtttggat

Amplification of the Enhancers	
Sall1-E1-Fwd	TCTGGAGAACAACACTCACACCC
Sall1-E1-Rev	GCAAGCCAGTAGATACCGCA
Sall1-E2-Fwd	ACTCTTTCACATGTGCCAAA
Sall1-E2-Rev	TCCAGCACAAGAAATCCTGC
Sall1-E3-Fwd	GAAAAAAAAAATCTTAGGTGC
Sall1-E3-Rev	GAGCAAACAACAGCCTTCCC
Sall1-E4-Fwd	GTTTGTTC AATTTTTAAATT
Sall1-E4-Rev	ACATTGGCCTAGAAGGTATC
Sall1-E5-Fwd	CAGGGGAAGGAAGGCAGGCT
Sall1-E5-Rev	GTGGGACCCCTGCCGGTGGC

PiggyBac Wt1-Sall1-BAC Cloning	
Sall1-CDS-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTTTGAGCCAGCATGTCGCGG
Sall1-CDS-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTTCTGGCAGCTTTAGCTTGTG
Neo-Rec-Fwd	TGGGTAAGGCAGTGATGACAGATCAAAGTAAAAGGTCTCACCCAGTCTACTCGACTGC ACGCGTTATATAG
Neo-Rec-Rev	TAAATAACCCCTCCTTTGTGTTCTTAACCCACTTAAATTTATTGCTTCATGTACCTGA CTGATGAAGTTC

Genotyping Sall1-BAC insertion into ES cells	
Sex-PCR-Fwd	CTGAAGCTTTTGGCTTTGAG
Sex-PCR-Rev	CCACTGCCAAATTCTTTGG
5'ITR-BAC-Fwd	gacgcatgcattcttgaat
5'ITR-BAC-Rev	atgcgtcattttgactcacg
3'ITR-BAC-Fwd	gaagaaatttgagttttgtttt
3'ITR-BAC-Rev	cgcatgtgtttatcggctt
bck-BAC-Fwd	GGCGGTGTTGATACAGCGGGTAA
bck-BAC-Rev	CCGGCGTTCGGTCGAAGAGTATC

MATERIAL AND METHODS

Animal models

Adult, infant or embryonic specimens of the Iberian mole *Talpa occidentalis* were used with annual permission from the Andalusian Environmental Council granted to Prof. Rafael Jiménez. The animals were captured alive in poplar groves plantations in Santa Fe, Chauchina and Fuentevaqueros (Granada province, southern Spain) using an efficient trapping system as described in a previous publication⁴⁷ and handled according to the guidelines and approval of the "Ethical Committee for Animal Experimentation" of the University of Granada.

Hedgehogs (*Atelerix albiventris*) were maintained in the LANE animal facility at the University of Geneva and were sampled under the experimentation permit GE24/33145 approved by the Geneva cantonal veterinary authorities, Switzerland.

Shrews (*Sorex araneus*) were trapped in wooden traps and euthanized with an isoflurane overdose followed by open-heart perfusion (see⁴⁸ for details) in Möggingen, Germany, under permit number 35-9185.81/G-11/21 to Dina Dechmann.

LacZ transgenic mice were carried out at Lawrence Berkeley National Laboratory (LBNL, CA, USA) was reviewed and approved by the LBNL Animal Welfare Committee. Transgenic mice were housed at the Animal Care Facility (the ACF) at LBNL. All transgenic experiments were performed in accordance with national laws and approved by the national and local regulatory authorities. Mice were monitored daily for food and water intake, and animals were inspected weekly by the Chair of the Animal Welfare and Research Committee and the head of the animal facility in consultation with the veterinary staff. The LBNL ACF is accredited by the American Association for the Accreditation of Laboratory Animal Care International (AAALAC).

The experiments for *Sall1* overexpression transgenic mice were performed as approved by LAGeSo Berlin under license numbers G0346/13 and G0247/13. Transgenic experiments were performed using mouse embryonic stem cells (mESCs) from a C57BL/6J or C57BL/6J-129 hybrid background. For RNA-seq and CHIP-seq experiments, gonads from wild-type CD1 mice were used.

Histological and immunostaining analyses

Gonads from adult animals, infants and embryos were fixed in 4% PFA and embedded in paraffin. The embedded samples were sectioned in 5µm thick slides and stained with hematoxylin-eosin according to standard protocols.

For protein spatio-temporal detection experiments, indirect immunofluorescence was used. In brief, sample slides were incubated overnight with the primary antibody at a dilution according to the manufacturer's instructions. Next, samples were incubated with specific Alexa secondary antibodies 488 and 568 together with DAPI for one hour at room temperature. Slides were then mounted in fluoromount-G solution (SouthernBiotech) and pictures were taken either with a Laser confocal Zeiss LSM700 or Zeiss Axiovert 200M microscope. Primary antibodies and working dilutions are listed here: mouse anti-SALL1 (abcam ab41974, dilution 1:100), goat anti-FOXL2 (abcam ab5096, dilution 1:200) and rabbit anti-SOX9 (Cell Signaling #82630, dilution 1:200).

RNA isolation and cDNA synthesis

Total RNA was extracted from adult ovaries and kidney from hedgehog (*Atelerix albiventris*) and shrew (*Sorex araneus*) using RNeasy Mini Kit (Quiagen, 74106) according to the manufacturer's instructions. In short, the tissues were homogenized in RTL buffer supplemented with β -Mercaptoethanol and applied to spin columns. Genomic DNA was removed using RNase-Free DNase Set (Quiagen, 79254). Eluted RNA quality and concentration were measured using NanoDrop 2000 UV spectrophotometer.

1 μ g RNA per sample was used for reverse transcription into cDNA using SuperScript IV First-Strand Synthesis System (Invitrogen, 18091050) according to the manufacturer's instructions. In short, random hexamer primers were annealed to template RNA and RNA was reverse transcribed into cDNA. Finally, RNA was removed using RNase H and RT reaction was used for RT-qPCR.

RT-qPCR *SALL1* and *FOXL2* mRNA levels were quantified by RT-qPCR for 2 biological replicates each in technical triplicates. RT-qPCRs were performed using 2X Blue S'Green qPCR Kit Separate Rox (Biozym, 331416) according to the manufacturer's instructions with 27.5 ng cDNA and 100 nM of each primer. All experiments were performed on QuantStudio 7 Flex system (Thermo Fisher).

Expression levels were normalized to *RPS9* mRNA. $2^{-\Delta\Delta Ct}$ method was used for analysis of relative *SALL1* and *FOXL2* expression levels. One-tailed t-test statistical tests were applied in these experiments.

ChIP sequencing

Gonads from E13.5 mouse embryos were fixed using 1% formaldehyde and subsequently snap-frozen and stored at -80°C . Chromatin immunoprecipitations were performed using the iDeal ChIP-seq Kit for Histones (Diagenode, Cat. No. C01010051) according to the manufacturer's

instructions. Briefly, whole fixed gonads were lysed and subsequently sonicated using a Bioruptor (45 cycles, 30 seconds on, 30 seconds off, at high power) in the provided buffers. 5 µg of sheared chromatin were then used per immunoprecipitation with 1 µg of the following specific histone antibodies: anti-H3K4me3 (Millipore, cat. No. 07-473), anti-H3K4me1 (Diagenode, cat. No. C15410037) and anti-H3K27ac (Diagenode, cat. No. C15410174). The samples were sequenced using Illumina HiSeq technology according to standard procedures. Mapping was performed with the STAR v2.6.1d software⁴¹ using settings to enforce unspliced read mapping (-alignEndsType EndToEnd --alignIntronMax 1 --outFilterMatchNminOverLread 0.94). Finally, deduplication was performed via bamUtil (version 1.0.14; option -rmDups, <https://github.com/statgen/bamUtil/releases>). Previous published ChIPseq data from mole developing gonads¹¹ were used to call putative enhancer regions.

Enhancer calling and conservation

Calling of putative enhancer regions was performed for mole and mouse via the software CRUP with replicates merged beforehand. CRUP software combines profiles from three histone marks, H3K4me3, H3K4me1 and H3K27ac, to define active enhancers. Enhancer regions with a distance ≤ 200 bp were merged. To reduce outlier effects in enhancer probability scores, a smoothing over 5 bins of 100bp was applied. In line with the original CRUP results, the probability of an enhancer region is defined as the, now smoothed, maximum score of the 100bp bins overlapping the enhancer. For the analysis of enhancer conservation, mole enhancer regions were lifted-over to the mouse genome (mm9). By definition, only those regions overlapping a conserved sequence block can be lifted, therefore depending on genome alignment settings. Here, we performed a sensitive pair-wise one-to-one genome alignment using LAST with automated training of optimal alignment parameters. In cases where an enhancer overlaps a conserved block partially, the respective non-conserved boundary is interpolated by the distance to the closest conserved block. Accordingly, the size of the lifted enhancer region in mm9 will be approximately the same as the one of the respective mole enhancer. Nevertheless, to exclude artefacts, lifting is only accepted if the ratio of mole enhancer length / lifted length is < 1.5 . We define an enhancer sequence as conserved if the enhancer could be lifted successfully. In addition, we define an enhancer as conserved in enhancer function, if the mole enhancer overlaps a mouse enhancer irrespective of tissue-specificity.

Transcriptomic analyses

For gene expression analysis, gonads from adult mice and embryos at E13.5, were dissected and RNA was extracted from these samples using the RNeasy Mini Kit (QIAGEN) according to the

manufacturer's instructions. For mole gonads previously published RNA-seq data was used⁷. The samples were sequenced using Illumina HiSeq technology according to standard procedures. Read mapping was performed with the STAR v2.6.1d software⁴⁹. Read counts were created using the R function "summarizeOverlaps" and normalized to RPKM based on the number of uniquely mapped reads. For the analysis of differential expression between samples, the DESeq2 tool was used with default settings⁵⁰.

Definition of female testis part specific regions

In order to prioritize enhancers by their potential relevance in testis part tissue, we first ranked enhancer regions by the difference in enhancer probability (score in testis part vs mean of scores in testis + ovary part). We defined the putative target genes of each enhancer as the gene with the closest transcription start site to the center of the enhancer region within the same TAD. Based on the differential expression analysis (testis part vs testis + ovary part), each target gene is ranked by specific expression in ovotestis (log₂ fold-change). Finally, enhancers are ranked jointly for functional importance in the female testis part by the mean rank of probability score and the rank of the putative target gene.

Transcription factor binding motif enrichment analysis

The five *SALL1* enhancer sequences from *Talpa occidentalis* were lifted-over to the genomes of mouse (UCSC:mm39) and *Sorex araneus* (UCSC:SorAra2.0) respectively based on pair-wise genome comparisons. In case of partial conservation enhancer boundaries were approximated given the sizes of the remaining non-conserved parts in *Talpa*.

For each enhancer sequence transcription factor (TF) binding affinities were computed via TRAP⁵¹ for all TransFac motifs (release TFP_2022.2). In case of TFs represented by multiple motifs only the one with the smallest p-value was kept. In addition, TFs not expressed in gonads (RPKM<3) were discarded. Finally, for each species TFs were ranked by the mean -log(p-value) across the group of five enhancers. To avoid artefacts introduced by non-significant binding affinities p-values < 0.05 were set to this cutoff. As a consequence, the computed mean should roughly correlate with the number of TF binding sites. Typically, only a subset of enhancers shares significant binding affinity for a specific TF (**Supplementary Table 2, column 6**) that shows a sufficient expression (**column 4**) in at least one gonadal tissue. Original affinity p-values for enhancers 1-5 are listed in the last column.

Hi-C

Previously published datasets from mole embryonic limb buds and adult ovotestes were used to inspect the *SALL1* regulatory domain⁷. Maps were visualized with the Juice box software⁵².

Mouse Hi-C was obtained from publicly available high-resolution datasets from neuronal progenitor cells (NPCs)²². Maps were visualized with the Juice box software⁵².

4C sequencing

Embryonic tissues were dissociated with trypsin, filtered through a cell strainer to obtain a single cell suspension and subsequently fixed in 2% formaldehyde. Mouse embryonic stem cells (mESCs) were detached from culture plates and fixed in the same way. Cells were counted and five million cells were snap-frozen and stored at -80°C until processing.

4C-seq libraries were prepared according to standard protocols⁵³. For the initial digestion, *Nla*III was used in *SALL1* experiments and *Bfal* was used in ITR-BAC ES cells. For the second digestion, *Dpn*II was used for all experiments. A total of 1.6 mg of each library was amplified by PCR for each viewpoint with primers listed in **Supplementary Table 4**. The libraries were sequenced using Illumina HiSeq technology according to standard procedures. Raw reads were pre-processed and mapped to the reference genome (ta10cc4) using BWA⁵⁴. Finally, reads were summarized and normalized by coverage (RPM) for each fragment generated by neighboring restriction enzyme sites. The viewpoint and its flanking fragments (1.5 kb upstream and downstream) were removed for data visualization and a window of 10 fragments was used to smoothen the data.

The mouse virtual 4C profile was derived from a genome-wide Hi-C map from NPCs¹⁷ by first extracting the intrachromosomal contact maps for the chromosomes of interest using Juicer tools v0.7.5⁵⁵ (KR normalized, MAPQ \geq 30, 5kb resolution). Afterwards, only map entries with at least one bin overlapping the viewpoint (chr8:89,044,162 (*Sall1*) on mm10) were used for the virtual 4C profile.

LacZ reporter assay in transgenic mice

LacZ transgenic mouse reporter assays were conducted as described⁵⁶. Briefly, enhancer sequences were amplified by PCR from mole genomic DNA using primers listed in **Supplementary Table 4**. PCR products were cloned into a vector containing a minimum

promoter, hsp68, in front of the LacZ gene. For microinjection into fertilized eggs, plasmid DNA was linearized with PacI and purified using Montage PCR filter units and Micropure EZ column (Millipore). For pronuclear injection of FVB embryos, DNA was diluted to a final concentration of 1.5-2 ng/ μ l and used in accordance with standard protocols approved by the Lawrence Berkeley National Laboratory. Embryos were harvested at embryonic day 13.5, dissected and fixed in 4% paraformaldehyde (PFA). Tissues were stained for 24 hours with freshly prepared staining solution, washed and post-fixed in 4% PFA.

BAC transgenesis for overexpression of Sall1

SALL1 coding sequence (CDS) was amplified from a vector containing the cDNA mouse sequence (Origen, cat. No. MC203471) with specific primers compatible with the attB gateway recombination system (Invitrogen). Through the gateway system, the generated product was introduced into a modified Wt1-BAC, containing piggyBac DNA transposon elements, as well as attL docking sites. The Wt1-BAC vector was kindly provided by Dr. Koopman and its further modification was performed according to their previously published method²⁶. After introduction of the *SALL1* minigene, a eukaryotic antibiotic resistance (dual Neomycin-Kanamycin cassette) was introduced into the BAC vector through recombineering for transfection into ES cells according to the protocol previously described⁵⁷. Primers are listed in **Supplementary Table 4**.

BAC transfection into female ES cells

Blastocysts from C57BL/6J mice were used to derive mouse embryonic stem cells (mESCs) by growing them with culture medium supplemented with leukemia inhibitory factor (LIF), as well as FGF/Erk and Gsk3 pathway inhibitors (2i). The derived mESCs were genotyped for sex and a female line was expanded through co-culture with mouse embryonic fibroblasts (MEFs) for further experiments.

Female mESCs were co-transfected with 3 μ g piggybac transposase and 500 ng of the modified Wt1-*SALL1*-piggyBac-Neo-BAC using Lipofectamine LTX (Invitrogen), as described in a previous publication⁵⁸. After Geneticin-G418 selection (250 μ g/ml) for 5 to 10 days, clones were picked and checked for successful BAC integration with 3 genotyping PCRs. A primer pair targeting each piggybac ITR (5'ITR and 3'ITR) was used as positive control, while a primer pair targeting the BAC vector was used as negative control to confirm integration mediated by transposition, instead of random insertion. Positive clones were expanded and additional genotyping was done by 4C-seq, to confirm genomic integrations site, as well as number of integrations, as described previously⁵³.

Gene ontology analyses

For Gene Ontology (GO) terms enrichment analysis PANTHER software was used⁵⁹, selecting all the common upregulated genes for the testis part of the ovotestes and in the *Sall1*-overexpressing mouse mutants. A total of 56 genes were evaluated. No significant enrichment was found for the downregulated genes.