

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 µm.

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:  $20 \mu m$ .

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 ± SD and p-values are indicated as \* $p \le 0.05$ , \*\* $p \le 0.01$ , \*\*\* $p \le 0.001$ .

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 ± SD and p-values are indicated as \* $p \le 0.05$ , \*\* $p \le 0.01$ , \*\*\* $p \le 0.001$ .



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

- A. Hi-C maps at high resolution from embryonic limbs with the corresponding TAD calling (black bars) underneath.
- B. 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.



В



#### 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 Gbrowse<sup>46</sup>.

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 Gbrowse<sup>46</sup>.

	tctggagaac	actcacaccc	ttcttatatt	agaatttctt	aggataatct	50	tctggagaac	actcacaccc	ttcttatatt	agaatttctt	aggataatct	50
	ctcaaattat	ttacacttga	attctggtct	cggggtgtat	ctctggggta	100	ctcaaattat	ttacacttga	attctggtct	cggggtgtat	ctctggggta	100
	acccatataa	gacagcccag	gcatctcact	gagataagat	gtgtttctgg	150	acccatataa	gacagcccag	gcatctcact	gagataagat	gtgtttctgg	150
	atgggagggg	cctcatgtgt	aaaaccaaaa	aggtgctaag	tgagcactta	200	atgggagggg	cctcatgtgt	aaaaccaaaa	aggtgctaag	tgagcactta	200
	ctacttgcct	gcctctgtgc	agtttatcgg	ccagatctca	cgataactcg	250	ctacttgcct	gcctctgtgc	agtttatcgg	ccagatctca	cgataactcg	250
	gggcatgcta	cggtagacac	aactgccacc	tctatttac	agataagaca	300	gggcatgcta	cggtagacac	aactgccacc	tctatttac	agataagaca	300
	aatgaaattc	aaaaagatta	atatggaccc	agattgttgg	ctgcagagtt	350	aatgaaattc	aaaaagatta	atatggaccc	agattgttgg	ctgcagagtt	350
	tgcctctgag	gttgctcctc	cctggtgccc	agcacctcgc	tcccttgtct	400	tgcctctgag	gttgctcctc	cctggtgccc	agcacctcgc	tcccttgtct	400
<b></b>	agagcatcat	taggccttcc	cactgcagaa	ccgggccttg	ctcctgggtg	450	agagcatcat	taggccttcc	cactgcagaa	ccgggccttg	ctcctgggtg	450
Т	gacacgctct	gggagctaga	cacgagtctg	ctcttcctgc	actgggtagg	500	gacacgctct	gggagctaga	cacgagtctg	ctcttcctgc	actgggtagg	500
L	gtgctaccag	ctgctggggc	caagattcat	gcttccaaga	aaagtgtgct	550	gtgctaccag	ctgctggggc	caagattcat	gcttccaaga	aaagtgtgct	550
ā	tttgaagaaa	aaaaaaaaaa	aaggaaggga	gagagaagag	gtaggaggga	600	tttgaagaaa	aaaAAAAAAA	AAGgaaGGGA	GAGAGAAGAg	gtaggaggga	600
×	gactaaatac	aaaaaaaaaa	aaaaaaaaaa	aaggactccg	ggagtettet	650	gactaaatac	aaaaaaaaaa	aaaaaaaaa	aaggactccg	GGAGtCTTCT	650
2	TGGAAAGCAA	aCageTTgTG	CAAAGCAACA	TTTTGATGAG	TTAAGTGCAA	700	TGGAAAGCAA	ACAGCTTgTG	CAAAGCAACA	TTTTGATGAG	TTAAGTGCAA	700
	ATGAATCAGG	AATAATGAGA	AACACaTGGG	TAAAETTTC	ССТААААТАТ СССТААААТАТ	750	ATGAATCAGG	AATAATGAGA	AACACATGGG	UTAAATTTTC mmcaccaaaa	СТТААААТАТ ПСПСПППППА А	/50
	COMPRESSOR	AAATTAAGAG	ttagaaaaat	TIGGGGAAAAA	atttattaa	800	CUTTA	AAATTAAGAG	GAAAAAGIGC	CCATCAACAT	AMMMCCAMCAC	800
ř	GITTagecaa	actadatta	atagatagaa	GCCACATACA	arcamaaaca	900	CCCCCAAAAT	GACTACCTTC	GTCCATCCC2	GCCACA+ACA	CTCATAAACa	900
<u> </u>	Cocceptaaat	gaetaggetg	agagatttta	GOCAGATAGA	ggagagaatt	950	CACACATAAA	CAC+CCACCa	agagatttta	GOCAGACAGA	ggagagaatt	900
2	taacaacaat	cagtggagga	toottootto	gattagecca	ggacagaatt	1000	taacaacaat	castotagaa	tagatagata	gattageta	ggacagaatt	1000
ш	caacaataatt	ttacttacta	tatatata	accaacacaa	ggatttagtg	1050	caacaacaac	ttacttacta	tatgtgtgtg	accaaca	GCACTCage	1050
_	gcactetgaa	tataacceta	ataaaaaaa	aatacagtcg	tttgacaage	1100	gcactctgaa	tataacceta	ataaaaaaa	aatacagtcg	tttgagaage	1100
	tatttactaa	aaacaccaaaa	tatatttaat	actttttaa	aaaaacatto	1150	tatttacta			acttttttaa	aaaaacatto	1150
	tcccaaatct	atattagtga	gacaaagcag	gaagatgtgt	taggattcca	1200	toccaaatot	atattagtga	gacaaagcag	gaagatgtgt	taggattcca	1200
	gagacttcca	acttotaget	ggcaataatt	gaagaegege	aacaaqqacq	1250	gagacttcca	acttctaget	ggcaataatt	gaagaegege	aacaaqqacq	1250
	gaacacagaa	agcaggaccc	ctctagccag	gcagaggggg	agttcccagg	1300	gaacacagaa	agcaggaccc	ctctagccag	gcagaggggg	ggttcccagg	1300
	agggcagagc	aagactttgc	tcctcagccc	taaactaaaa	taaaatata	1350	agggcagagc	aagactttgc	tcctcagccc	tagactagaa	taaaatata	1350
	tgaaggcatc	cctttccaaa	cccacaggga	gctacagcgc	tttccactgg	1400	tgaaggcatc	cctttccaaa	cccacaggga	gctacagcgc	tttccactgg	1400
	ttcccattca	cgtgctcagg	aacaatgatc	ctgattccat	ttatattttc	1450	ttcccattca	cgtgctcagg	aacaatgatc	ctgattccat	ttatattttc	1450
	aaccagcgta	actttcagct	aataaaaaaa	tggagtttct	ggaaattggt	1500	aaccagcgta	actttcagct	aataaaaaaa	tggagtttct	ggaaattggt	1500
	gcttattttt	aaaqaaqcaa	tagtgcggta	tctactggct	tac		gcttattttt	aaaqaaqcaa	tagtgcggta	tctactggct	tac	
	<b>,</b>	j j			- ,-		J	j j			- 5-	
	TTGGATTTTA	ACCAAGTGAG	GGGTTATTT	GCCTTTATA	тааассттаа	50	TTCCATTTA	ACCAAGTGAG	GGGTTATTT	GCCTTTATA	тааассттаа	50
	AAGGCATATA	Сататстатс	ATTTACCTAT	CCATTTATA	TARAGETTAA	100	AACCCATATA	Сататстатс	ATTACCTAT	CCATTTATA	+CTTTGAGTC	100
	+TTTGAATAT	TAATTAGCAG	GAACTTCAGA	AGGCAAcaGG	GACACATTTT	150	TT++GAATAT	TAATTAGCAG	GAACTTCAGA	AGGCAACAGG	GACACATTTT	150
	GTTTccaata	tctcgcaatt	tagaccaatt	cactcoatga	atagaactac	200	GTTTCCAATA	тстсасаатт	TGGGccaatt	cactocatga	atagaactac	200
-	tcatacttac	acaacaaaat	ctattttact	gcaatggtgg	attacaccca	250	teatacttac	acaacaa	CTATTTTact	CCAATCGTCC		250
Ν	gatggctccc	caaacaaaaa	getecagea	aagaaatcot	accttattat	300	GATGGCTCCC	cGGGCgaGGg	GCTCcaGCAc	aaGAAATCCT	GCCTTGTTGT	300
_	gacaggcgga	gaaaaggaat	agagacctag	agaccaatca	ggagctggat	350	GACAGGCGGA	GAAAaGGAAt	agagacctag	agaccaatca	ggagctggat	350
<u> </u>	ctggccggga	tctatatcta	cttcatggaa	aaactgccga	agctatctcg	400	ctggccggga	tctgtgtctg	cttcatggaa	aaactgccga	agctatctcg	400
Ð	agacttacat	ctttggaaac	gtcagcaggg	gttggggaga	tctacttctg	450	agacttacat	ctttggaaac	gtcagcaggg	gttggggaga	tctacttctg	450
Ū	ccccagacga	ataattetae	tgataattat	ccctqttccc	aactgttgac	500	ccccagacga	ataattetae	tgATAATTAT	CCCTGtTCCC	aacTGTTGac	500
ž	tacgaagggg	ccctctgtaa	aggcccctcc	aggtgtcaca	cataggagac	550	tacgaagggg	ccctctgtaa	aggecectee	aggtgtcaca	cgtgggagac	550
	qqqqtqtqcq	cacatatata	agtatatata	cgtgtcatgt	gtgtcacgtg	600	aggatataca	cacatatata	aatatatata	cqtqtcatqt	gtgtcacgtg	600
σ	tgcccatcct	cctccacttg	ggacattccc	tgagagcctt	ggccgcacgg	650	tgcccatcct	cctccacttg	ggacattccc	tgagagcett	ggccgcacgg	650
_	gagggagcag	caacagcaag	gcctcgagat	tgctttcact	ttcagtggcg	700	qaqqqaqcaq	caacagcaag	gcctcgagat	tgctttcact	ttcagtggcg	700
7	cctgagcacg	ctgageteta	ggtgacgttc	tcccccagcg	tggggatttc	750	cctgagcacg	ctgageteta	ggtgacgttc	tcccccagcg	tggggatttc	750
<u> </u>	ctgccactga	ctcgctgggc	acagaagtgc	gtccaggctg	gtgtgcagcg	800	ctgccactga	ctcgctgggc	acagaagtgc	gtccaggctg	gtgtgcagcg	800
ш	cagcacattc	ttatgcagaa	atgcaagcac	actaaggttt	gcttcttttg	850	cagcacattc	ttatgcagaa	atgcaagcac	actaaggttt	gcttcttttg	850
	ttttggcaca	tgtgaaagag	tttctacaaa	actgatttta	cagagcaata	900	ttttggcaca	tgtgaaagag	ttTCTACAAA	ACTGATTTta	CAGAGCAaTa	900
	agtcacctcc	aagggtgtcg	ccagccattg	aaagtgcacc	acagaggett	950	AGTCACCTCC	Aagggtgtcg	ccagCCATTG	AAAGTGCACC	aCaGAGGCTT	950
	ttttatgcgg	taaatgcact	gggaaggtag	tgtgtaatgg	cccagcgcct	1000	TtttAtgcgg	TAAATgCAcT	GGGAAGGTAG	tgtgtaatgg	CCCAGcGCCT	1000
	ggactccagC	TCCATTTAGC	cCAgACAGCC	TTGTTTTTAC	TGCCCCTTAG	1050	GgACTCCAGc	TCCATTTAGC	ccagACAGCC	TTGTTTTTAC	TGCCCCTTAG	1050
	TGCCTCACTT	TTAAAGTGAC	AGGAAAcCAA	TTTCTGATTA	CAAATCAATG	1100	TGCCTcaCTT	TTAAAGTGAC	AGGAAACCAA	TTTCTGATTa	CAAATCAATG	1100
	G						G					
	ggtttccctc	ctgctcactt	gcagccttcg	cttttaagcc	tggtttaaac	50	ggtttccctc	ctgctcactt	gcagccttcg	cttttaagcc	tggtttaaac	50
	acacccgccC	TCCCTTcttt	ttaTTTAGAA	AGAAAAaaaa	aTCTTAGGTG	100	acacccgcCC	TCCCTTCTTT	TTATTTAGAA	AGAAAaaaaa	ATCTTAGGTG	100
	cTCtgtTTAC	TTACaTTTCT	CAGCACTTCT	TTGTCctCCA	AAtCactcTa	150	CTCTGTTTAC	TTACATTTCT	CAGCACTTCT	TTgtcctcca	aatcactcta	150
	gCTGaTgaTA	TGGAAGtTtt	AtggGctATT	Ttatcagaaa	AgGCAGACCT	200	gctgatgata	tggaagtttt	atgggctatt	ttatcagaaa	agg <mark>CAGACCT</mark>	200
	tTGCTAGCTG	<b>AC</b> ctgggcag	ttggaacagg	tttaagtgtg	catggaaacc	250	TTGCTaGctG	ACcTGgGCaG	TTgGaAcaGG	tTTAAGTGTG	CATGGAAACC	250
	cctgatttca	cccgaaaagc	aatacttggc	ataggttgac	atctagagcc	300	CCTGATTTCA	CCCGAAAAGC	AATactTGGC	ATAgGtTgAC	ATCTAGAGCC	300
	tgaaaACACA	GCCAGtgaGA	GTcGCgCAGA	CTAAAGGAAA	TCACTaCTTA	350	TGAAAACACA	GCCAGTgAGA	GTCGCGCAGA	CTAAAGGAAA	TCACTACTTA	350
	ATTCCACAAA	CATTtgtttg	ggt <mark>CCAGGAA</mark>	<b>TGGCACA</b> aga	aacttgaagc	400	ATTCCACAAA	CATTTGtTTg	GGTCCAGGAA	tGGCACaAGA	AActtgaagc	400
	ctctggctgc	ggcgccggcc	gggcacgcga	ggggccatag	cgcgcggggc	450	ctctggctgc	ggcgccggcc	gggcacgcga	ggggccatag	cgcgcggggc	450
	agcgccacct	gccggccggc	ctggcacccc	gagagggccc	cgcggcccgg	500	agcgccacct	gccggccggc	ctggcacccc	gagagggccc	cgcggcccgg	500
	cgcccctgcc	atccgctggc	ttgggttccc	gacctggtcc	cctggggctc	550	cgcccctgcc	atccgctggc	ttgggttccc	gacctggtcc	cctggggctc	550
m	cctgctgctg	gggagctctg	actgcaccag	agggaagaga	gggcggctgt	600	cctgctgctg	gggagctctg	actgcaccag	agggaagaga	gggcggctgt	600
-	gctggagaga	gacagccctg	cccgggaaac	cctcccaagg	cctgtttcac	650	gctggagaga	gacagccctg	cccgggaaac	cctcccaagg	cctgtttcAC	650
<u> </u>	atgtagaaga	gggaagatta	atacctctgg	agccggcccg	gccgtcctgc	700	ATGTAgAaGa	gggaagatta	ATACCTCTGG	AgCcGgCCCg	gccg <mark>TC</mark> ct <mark>GC</mark>	700
Ð	caagggcaca	ggtttccagc	actagaggtg	aagcggagtt	tatttgcatc	750	CaaGGGCACA	gGTTTcCAGC	ActAgAGGTG	AAGCgGAGTT	TATTTGCATC	750
U	cccagtaaag	ggttgttttc	ccttctttct	tgcgggctat	aaaacactgc	800	Cccagtaaag	ggttgttttc	ccttctttct	tgcgggctat	aaaacactgc	800
ř	ctcttaacca	ccttgtcttg	caccgggttc	tgggggctgc	ttttaggggt	850	ctcttaacca	ccttgtcttg	caccgggttc	tgggggctgc	ttttaggggt	850
=	gcagtgggcg	ccatcctgga	acacccgccg	tcgggggggc	caggttaatt	900	gcagtgggcg	ccatcctgga	acacccgccg	tcgggggggc	caggttaatt	900
σ	cgtctcctgt	ctgtctgcct	tggggagggg	ggcagcgggg	gtatagggag	950	cgtctcctgt	ctgtctgcct	tggggagggg	ggcagcgggg	gtatagggag	950
2	gaggcagcag	tgcttttgga	gacactgggg	cgctgtctgt	ccacaatagg	1000	gaggcagcag	tgcttttgga	gacactgggg	cgctgtctgt	ccacaatagg	1000
-	gtccaggaat	cagtgcaggg	aggcccccac	gagggcagaa	gtcggggacg	1050	gtccaggaat	cagtgcaggg	aggcccccac	gagggcagaa	gtcggggacg	1050
	ggggaaatgc	catcagetae	ttetgegtae	ttccaccctg	cettgtccgg	1150	ggggaaatgc	catcagctac	ttctgcgtac	ttccaccctg	ccttgtccgg	1150
ш	ggtcccagat	ttcagggggc	tcatcgctga	taaactgcaa	acggcgcaag	1200	ggtcccagat	ttcagggggc	tcatcgctga	taaactgcaa	acggcgcaag	1150
	ctaaatggac	caaggaggcg	gecattacte	LCCgggggagg	yeetcacggg	1200	ctaaatggac	caaggaggcg	gccattactc	ttcggggagg	gcctcacggg	1250
	aacgcaggat	cgagtgttat	ttgcaactca	trattagge	LgtgCtggga	1200	aacgcaggat	cgagtgttat	etgeaactea	tratteaagge	cgtgctggga	1200
	ttoottoo	caagtgcctt	tototototot	cgatttccat	tatagagtge	1350	catecactec	caagtgeett	tgcaagetg	cgatttccat	tatagatta	1300
		accocccttt	ctagaagaga	agaagactee	COLCECTIC	1400	cottocag	acceccett	ctoccetetg	getetgtgee	CELECCETTC	1400
	cattagetet			auaduaCtCa	yacccagtcc	1400	tagattact	tangaget	ctggaagggg	ayaayaCtCa	yacccagtee	1450
	ccttcgatct	tanaa	ataattaa	atoot	agaaactet	1 /1 10 /1		- G > G A G G G G G G G G G G G G G G G G			aaaa	A 44 Min 2 M
	ccttcgatct tggcttccac	tgaccggtgg	ctggttcaga	atcctgccgt	gggcactctg	1450	cagtococc	aggaggtatata	accagagaga	atcetgeegt	gggcactctg	1450
	ccttcgatct tggcttccac cagtcccggg	tgaccggtgg aggagctgtg	ctggttcaga gccggggcga	atcctgccgt ggtcagctgg	gggcactctg cactgcctct	1450 1500 1550	cagtcccggg	aggagetgtg	gccgggggcga	ggtcagctgg	gggcactctg cactgcctct	1450
	ccttcgatct tggcttccac cagtcccggg gcgcacgaga	tgaccggtgg aggagctgtg ggaccagaga	ctggttcaga gccggggggga tgatgcttct	atcctgccgt ggtcagctgg tcaccacgct	gggcactctg cactgcctct ctccctaggc gacagaacta	1450 1500 1550 1600	cagtcccggg gcgcacgaga	aggagetgtg ggaccagaga	gccggggggga tgatgcttct	ggtcagctgg tcaccacgct	gggcactctg cactgcctct ctccctaggc	1500 1550 1600
	ccttcgatct tggcttccac cagtcccggg gcgcacgaga caaccagcat	tgaccggtgg aggagetgtg ggaccagaga ggegetgece	ctggttcaga gccggggcga tgatgcttct atgaacaaca	atcctgccgt ggtcagctgg tcaccacgct tagcttgggg	gggcactctg cactgcctct ctccctaggc gacagaacta aggtgtgggt	1450 1500 1550 1600	cagtcccggg gcgcacgaga caaccagcat	aggagctgtg ggaccagaga ggcgctgccc	gccggggcga tgatgcttct atgaacaaca	ggtcagctgg tcaccacgct tagcttgggg	gggcactctg cactgcctct ctccctaggc gacagaacta	1450 1500 1550 1600
	cettegatet tggettecae cagteceggg gegeaegaga caaceageat catgggeeet	tgaccggtgg aggagctgtg ggaccagaga ggcgctgccc caagaggcct gggaaggctg	ctggttcaga gccggggcga tgatgcttct atgaacaaca ggcccccaag	atcctgccgt ggtcagctgg tcaccacgct tagcttgggg gggcccgtgg aggggctgct	gggcactctg cactgcctct ctccctaggc gacagaacta aggtgtgggt gcccgactag	1450 1500 1550 1600 1650 1700	cagtcccggg gcgcacgaga caaccagcat catgggccct	aggagetgtg ggaceagaga ggegetgece caagaggeet	gccggggggga tgatgcttct atgaacaaca ggcccccaag	ggtcagctgg tcaccacgct tagcttgggg gggcccgtgg agggctgct	gggcactctg cactgcctct ctccctaggc gacagaacta aggtgtgggt gcccgactgg	1450 1500 1550 1600 1650 1700
	cettegatet tggettecae cagteceggg gegeaegaga caaceageat catgggeeet aggagagtet	tgaccggtgg aggagctgtg ggaccagaga ggcgctgccc caagaggcct gggaaggctg tccacaagtg	ctggttcaga gccggggcga tgatgcttct atgaacaaca ggcccccaag ttgtttgctc ttcctccagc	atcctgccgt ggtcagctgg tcaccacgct tagcttgggg gggcccgtgg aggggctgct gtgacctgte	gggcactctg cactgcctct ctccctaggc gacagaacta aggtgtgggt gcccgactgg c	1450 1500 1550 1600 1650 1700	cagtcccggg gcgcacgaga caaccagcat catgggccct ctggccccta aggagagtct	aggagctgtg ggaccagaga ggcgctgccc caagaggcct gggaaggctg tccacaagtg	gccggggcga tgatgcttct atgaacaaca ggcccccaag ttgtttgctc ttcctccage	ggtcagctgg tcaccacgct tagcttgggg gggcccgtgg aggggctgct gtgacctatc	gggcactctg cactgcctct ctccctaggc gacagaacta aggtgtgggt gcccgactgg c	1450 1500 1550 1600 1650 1700

#### Against mouse

#### 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.

### Against shrew

- C
C
- L
4
-
<u> </u>
6
C
-
6
6
<u> </u>
_
6
C
2
2
nment • Sun
Supt - Sup
lonment • Sun
Suprant • Supr
alonment • Sun
ialonment • Sun
Valonment • Sun
Suparant Sup
evelopment • Sup
levelonment - Sun
Cupanent . Sup
Development • Sup

Against human

#### AGCTAGACTT GAATTTATTC AGCACUTATT TTAAAAATAA TTTCCCCCGAA TTATAAAAAT AATGTTTGTT CAATTTTTTAA ATTAGUTGCC AAUGTATAAA ATTagggatt ttacataaag ttgtcaATTT CAGGTTTCTT TCCAAAACTT AGAATTCTgG AACcagttca acTCATTCCC TCAACcaaca tgtgaccaaa 50 100 150 4 200 ggtagtactg gctgctcatg ggcggggcg tgcattttc ggttcctgg agtctcccTC ACTCCCTga AtCcCTTAGA TtTAGTcCAC ATCTCTCATT tAgATTGTCT GCTGGAAAct ataaaacttt gaatttacaa cccctgactc 250 Enhancer 300 350 tcagccgtat ttttttttt taaagaaaat aaacatttca agttatttca 400 tcagcagta tutttitt taaagaaat aacattica agtiattoa tcaggaagaa gggacattig gggaaagtica actataggca ggaaaagti acagtigott tatgTATTT TCCATCtCaG ATTTTCtott TCTCTTTAG CCCAGiggag caaaaacaaa aagagctgca gataaatgat gocttagga cTCATAAAGT TcAAaGTAAG CAATTATAct tcacttaca agatotatag caaaataaat tgcaatggat tgaatagta aaatattat aaaggcatg 450 500 550 600 650 700 taacttaga tcatcotta aaggagcag tagotacaga gaaggcagag aaggtocact tagtgcaatt ctgtgtcatc tcagtattc agcaatttot taaaagtcat gaaaaaagtt taaaaatact tootgttaaa ggatacottc 750 800 850 taggccaatg tcatcagtga aattaaccaa gaacaaagag aacgagaacc 900 agaaaaggaa tgtccccttc taggaaacga gggtccacag gag

#### Against shrew

agctacactt	gaatttattc	agcacttatt	ttaaaaataa	tttcccccaa	50
ttataaaaat	aatgtttgtt	caatttttaa	attagttgcc	aatgtataaa	100
attagggatt	ttacataaag	ttgtcaattt	caggtttctt	tccaaaactt	150
agaattctgg	aaccagttca	actcattccc	tcaaccaaca	tgtgaccaaa	200
ggtagtactg	gctgctcatg	ggcgaggccg	tgcattttc	agtttcctgg	250
agtctccctc	actccctgaa	atcccttaga	tttagtccac	atctctcatt	300
tagattgtct	gctggaaact	ataaaacttt	gaatttacaa	cccctgactc	350
tcagccgtat	tttttttt	taaagaaaat	aaacatttca	agttatttca	400
tcaggaagaa	gggacatttg	gggaagttca	actataggca	ggaaaaagtt	450
acagttgctt	tatgtatttt	tccatctcag	attttctctt	tctcttttag	500
cccagaggag	caaaaacaaa	aagagc <mark>TGCA</mark>	GATAAATGaT	GCTCTTAGCA	550
CTCATAAAgt	TCAAAGTaAG	CAATtAtACT	TCACTTTACA	AGATCtATAg	600
CAAAAtAAAT	TGcaATGGAt	TGAATAGTTA	AAATATTTAT	AAAGGCAtAG	650
<b>g</b> GAAAAGACC	AGAAAACAgA	CACACAgACa	cacacacaCA	CACTTtTGTC	700
TTAACTTaga	tcatccttaa	aagggagcag	tagctacaga	gaaggcagag	750
aaggtccact	tagtgcaatt	ctgtgtcatc	tcagtatttc	agcaatttct	800
taaaagtcat	gaaaaaagtt	taaaatact	tcctgttaaa	ggataccttc	850
taggccaatg	tcatcagtga	aattaaccaa	gaacaaagag	aacgagaacc	900
agaaaaggaa	tgtccccttc	taggaaacga	gggtccacag	gag	

#### Against mouse

	ggagcgatga	tgtgagaggg	agagagaggg	gtgggatgca	atctctgaaa	50
	acgctcatca	acatatttcc	agtctgcaaa	aggacaggtc	acgctggagg	100
	aacacttgtg	gaagactctc	ctccagtcgg	gcagcagccc	ctgagcaaac	150
	aacagccttc	cctaggggcc	agacccacac	ctccacgggc	cccttggggg	200
	ccaggcctct	tgagggccca	tgtagttctg	tcccccaagc	tatgttgttc	250
	atgggcagcg	ccatgctggt	tggcctaggg	agagcgtggt	gaagaagcat	300
	catctctggt	cctctcgtgc	gcagaggcag	tgccagctga	cctcgccccg	350
	gccacagctc	ctcccgggac	tgcagagtgc	ccacggcagg	attctgaacc	400
	agccaccggt	cagtggaagc	caggactggg	tctgagtctt	ctccccttcc	450
	agaggtgaaa	agagatcgaa	gggaagggga	gaggcacaga	gccagagaga	500
	gaaaaggggg	gtctggaagg	aagcactgca	ggatggaaat	cacagcttgc	550
	aaaaggcact	tgggagtgga	tgtcccagca	cagccttgaa	aatgagttgc	600
L)	agataacact	cgatcctgcg	ttcccgtgag	gccctccccg	aagagtaatg	650
	gccgcctcct	tggtccattt	agcttgcgcc	gtttgcagtt	tatcagcgat	700
	gagccccctg	aaatctggga	ccccggacaa	ggcagggtgg	aagtacgcag	750
θ	aagtagctga	tggcatttcc	cccgtccccg	acttctgccc	tcgtgggggc	800
υ	ctccctgcac	tgattcctgg	accctattgt	ggacagacag	cgccccagtg	850
Ē	tctccaaaag	cactgctgcc	tcctccctat	acccccgctg	ccccctccc	900
=	caaggcagac	agacaggaga	cgaattaacc	tggcccccc	gacggcgggt	950
σ	gttccaggat	ggcgcccact	gcacccctaa	aagcagcccc	cagaacccgg	1000
<b>_</b>	tgcaagacaa	ggtggttaag	aggcagtgtt	ttatagcccg	caagaaagaa	1050
-	gggaaaacaa	ccctttactg	gggatgcaaa	taaactccgc	ttcacctcta	1100
<u> </u>	gtgctggaaa	cctgtgccct	tggcaggacg	gccgggccgg	ctccagaggt	1150
ш	attaatcttc	cctcttctac	atgtgaaaca	ggccttggga	gggtttcccg	1200
	ggcagggctg	tctctctcca	gcacagccgc	cctctcttcc	ctctggtgca	1250
	gtcagagctc	cccagcagca	gggagcccca	ggggaccagg	tcgggaaccc	1300
	aagccagcgg	atggcagggg	cgccgggccg	cggggccctc	tcggggtgcc	1350
	aggccggccg	gcaggtggcg	ctgccccgcg	cgctatggcc	cctcgcgtgc	1400
	ccggccggcg	ccgcagccag	aggcttcaag	tttctTGTGC	CATTCCTGGa	1450
	cccaaacaAA	TGTTTGTGGA	ATTAAGtAGT	GATTTCCTTT	AGTCTGcGCg	1500
	ACTCtcaCTG	GCTGTGTttt	caggctctag	atgtcaacct	atgccaagta	1550
	ttgcttttcg	ggtgaaatca	ggggtttcca	tgcacactta	aacctgttcc	1600
	aactgcccag	GTCAGCTAGC	AaAGGTCTGC	CTTTTCTGAT	aAAATagCcc	1650
	aTaaAaCTTC	CATAtcAtCA	<pre>GctAgagtGa</pre>	TTTGGagGAC	AAAGAAGTGC	1700
	TGAGAAAtGT	AAGTAAacaG	AgCACCTAAG	ATTTTttTT	CTttctaaat	1750
	aaaaagaagg	gagggcgggt	gtgtttaaac	caggcttaaa	agcgaaggct	1800
	gcaagtgagc	aggagggaaa				

#### Against shrew

ggagcgatga	tgtgagaggg	agagagaggg	gtgggatgca	atctctgaaa	50
acgctcatca	acatatttcc	agtctgcaaa	aggacaggtc	acgctggagg	100
aacacttgtg	gaagactctc	ctccagtcgg	gcagcagccc	ctgagcaaac	150
aacagccttc	cctaggggcc	agacccacac	ctccacgggc	cccttggggg	200
ccaggcctct	tgagggccca	tgtagttctg	tcccccaagc	tatgttgttc	250
atgggcagcg	ccatgctggt	tggcctaggg	agagcgtggt	gaagaagcat	300
catctctggt	cctctcgtgc	gcagaggcag	tgccagctga	cctcgccccg	350
gccacagctc	ctcccgggac	tgcagagtgc	ccacggcagg	attctgaacc	400
agccaccggt	cagtggaagc	caggactggg	tctgagtctt	ctccccttcc	450
agaggtgaaa	agagatcgaa	gggaagggga	gaggcacaga	gccagagaga	500
gaaaaggggg	gtctggaagg	aagcactgca	ggatggaaat	cacagcttgc	550
aaaaggcact	tgggagtgga	tgtcccagca	cagccttgaa	aatgagttgc	600
agataacact	cgatcctgcg	ttcccgtgag	gccctccccg	aagagtaatg	650
gccgcctcct	tggtccattt	agcttgcgcc	gtttgcagtt	tatcagcgat	700
gagccccctg	aaatctggga	ccccggacaa	ggcagggtgg	aagtacgcag	750
aagtagctga	tggcatttcc	cccgtccccg	acttctgccc	tcgtgggggc	800
ctccctgcac	tgattcctgg	accctattgt	ggacagacag	cgccccagtg	850
tctccaaaag	cactgctgcc	tcctccctat	acccccgctg	ccccctccc	900
caaggcagac	agacaggaga	cgaattaacc	tggcccccc	gacggcgggt	950
gttccaggat	ggcgcccact	gcacccctaa	aagcagcccc	cagaacccgg	1000
tgcaagacaa	ggtggttaag	aggcagtgtt	ttatagcccg	caagaaagaa	1050
gggaaaacaa	ccctttactg	gGGATGCAAA	TAAACTCcGC	TTCACCTcTa	1100
gTGCTGgAAA	CcTGTGCCCt	t <mark>GGC</mark> ag <mark>GA</mark> cg	gcc <mark>GGG</mark> cCgG	cTCCAGAgGT	1150
ATTAATCTTC	CctCtTcTAC	<b>ATGT</b> gaaaca	ggccttggga	gggtttcccg	1200
ggcagggctg	tctctctcca	gcacagccgc	cctctcttcc	ctctggtgca	1250
gtcagagctc	cccagcagca	gggagcccca	ggggaccagg	tcgggaaccc	1300
aagccagcgg	atggcagggg	cgccgggccg	cggggccctc	tcggggtgcc	1350
aggccggccg	gcaggtggcg	ctgccccgcg	cgctatggcc	cctcgcgtgc	1400
ccggccggcg	ccgcagccag	aggcttcaag	TTTCTtGTGC	Cattcctgga	1450
CCCAAaCAAA	TGTTTgTGGA	ATTAAGTAGT	GATTTCCTTT	AGTCTGCGCG	1500
ACTCTCACTG	GCTGTGTTTT	CAgGCTCTAG	ATGTCAaCCT	ATGCCAagtA	1550
TTGCTTTTCg	GGTGAAATCA	gGGGTTTCCA	TGCACACTTA	AaCCtgTtCc	1600
AACtGCcCAg	GTCagCtAGC	AAAGGTCTGC	cttttctgat	aaaatagccc	1650
ataaaacttc	catatcatca	gctagagtga	tttggaggac	AAAGAAGTGC	1700
TGAGAaATGT	AAGTAAACAG	AGCACCTAAG	ATtttttTTT	CTTTCTAAAT	1750
AAAAAGAAGG	GAGGgcgggt	gtgtttaaac	caggcttaaa	agcgaaggct	1800
gcaagtgagc	aggagggaaa				

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

The conserved nucleotides are highlighted in blue and capitalized. Light blues 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.





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:  $1000 \mu m$ , white scale bars:  $100 \mu m$ .



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  $\mu$ m, white scale bars: 100  $\mu$ m.



#### 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  $\mu$ m, white scale bars: 100  $\mu$ m.



#### 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  $\mu$ m, white scale bars: 100  $\mu$ m.



#### 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  $\mu$ m, white scale bars: 100  $\mu$ m.



### 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  $\mu$ 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  $*p \le 0.05$ ,  $**p \le 0.01$ ,  $***p \le 0.001$ .



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

Hematoxylin-eosin stanning 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:  $200 \mu m$ .

А	Mus musculus (REF)		<u>u</u>	<u>ipload_1</u> (▼ <u>Hierar</u>	chy	_ NEW! (?)	
GO biological process complete	<u>#</u>	<u>#</u>	expected	Fold Enrichment	<u>+/-</u>	raw P value	<u>FDR</u>
metanephric nephron morphogenesis	<u>18</u>	<u>3</u>	.04	67.89	+	1.79E-05	4.03E-02
•metanephros morphogenesis	<u>24</u>	<u>3</u>	.06	50.92	+	3.89E-05	4.09E-02
<u> </u>	<u>3306</u>	<u>20</u>	8.12	2.46	+	6.24E-05	4.68E-02
<mark> </mark>	<u>3813</u>	<u>25</u>	9.36	2.67	+	8.56E-07	1.35E-02
•multicellular organism development	<u>4559</u>	<u>26</u>	11.19	2.32	+	1.01E-05	2.64E-02
<u> </u>	<u>368</u>	<u>7</u>	.90	7.75	+	3.46E-05	4.20E-02
<mark> </mark>	<u>1057</u>	<u>11</u>	2.59	4.24	+	4.59E-05	4.02E-02
4anatomical structure morphogenesis	<u>2337</u>	<u>19</u>	5.74	3.31	+	1.50E-06	1.18E-02
<u> </u>	<u>136</u>	<u>5</u>	.33	14.98	+	2.45E-05	3.51E-02
positive regulation of branching involved in ureteric bud morphogenesis	<u>23</u>	<u>3</u>	.06	53.13	+	3.47E-05	3.90E-02
regulation of branching involved in ureteric bud morphogenesis	<u>25</u>	<u>3</u>	.06	48.88	+	4.35E-05	4.04E-02
regulation of morphogenesis of a branching structure	<u>66</u>	<u>4</u>	.16	24.69	+	2.61E-05	3.43E-02
regulation of multicellular organismal process	<u>2989</u>	<u>19</u>	7.34	2.59	+	5.28E-05	4.38E-02
ventricular septum development	<u>82</u>	<u>4</u>	.20	19.87	+	5.88E-05	4.63E-02
<u> </u>	<u>145</u>	<u>6</u>	.36	16.86	+	1.84E-06	9.66E-03
<u> </u>	<u>192</u>	<u>6</u>	.47	12.73	+	8.77E-06	2.76E-02
developmental growth involved in morphogenesis	<u>153</u>	<u>5</u>	.38	13.31	+	4.22E-05	4.16E-02
anatomical structure formation involved in morphogenesis	<u>980</u>	<u>11</u>	2.41	4.57	+	2.32E-05	3.65E-02
negative regulation of multicellular organismal process	<u>1154</u>	<u>12</u>	2.83	4.24	+	1.96E-05	3.87E-02
epithelium development	<u>1161</u>	<u>12</u>	2.85	4.21	+	2.08E-05	3.65E-02
<u> •tissue development</u>	<u>1785</u>	<u>16</u>	4.38	3.65	+	3.88E-06	1.53E-02
B <u>Mus musculus</u> (REF) <u>up</u>	<u>load_1</u> (▼ <u>Hierarchy</u> N	IEW	(?)				
GO cellular component complete # # expected F	Fold Enrichment +/- ra	w P	value F	-DR			

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.

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



Gene	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
Lhx1	76,0142476	0,416247256	0,45549231	0,91384036	0,36080074	0,99999929
Fgf9	34,7898814	-0,34199249	0,64606248	-0,5293489	0,59656341	0,99999929
Cxcr4	269,917075	0,363473378	0,24191365	1,50249224	0,13297003	0,99999929
Cxcl12	867,206005	-0,400398642	0,1872609	-2,138186	0,03250165	0,99999929
Spry2	246,127611	0,017530695	0,24731929	0,07088285	0,943491	0,99999929
Mmp14	2980,353	-0,004484855	0,11824949	-0,0379271	0,96974584	0,99999929

#### 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  $\mu$ 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.

Click here to download Table S1

## Table S2. Ranking of transcription factors by significance of binding affinity to the five *SALL1* enhancer sequences.

Click here to download Table S2

**Table S3.** Differential gene expression between Sall1-overexpressing mutant and wildtype ovaries.

Click here to download Table S3

#### Table S4. Primer list.

	RT-qPCRs
Hedgehog-qPCR-Sall1-Fwd	GAAGCAAGCGAAGCCTCAAC
Hedgehog-qPCR-Sall1-Rev	TGCTCTTAGTGGGGCGATTT
Hedgehog-qPCR-Foxl2-Fwd	CAGAAGCCGCCCTATTCGT
Hedgehog-qPCR-Foxl2-Rev	GGGAACTTGGCGATGATGT
Hedgehog-qPCR-Rps9-Fwd	GCCAAGTCCATCCACCAC
Hedgehog-qPCR-Rps9-Rev	CCAGGCGGACAATGAAGG
Shrew-qPCR-Sall1-Fwd	AGAGCGTTCACAACAAAGG
Shrew-qPCR-Sall1-Rev	TGGGGCCATCCACAGAGA
Shrew-qPCR-Foxl2-Fwd	CATCGCCAAGTTCCCCTTCT
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	TCTGGAGAACACTCACACCC		
Sall1-E1-Rev	GCAAGCCAGTAGATACCGCA		
Sall1-E2-Fwd	ACTCTTTCACATGTGCCAAA		
Sall1-E2-Rev	TCCAGCACAAGAAATCCTGC		
Sall1-E3-Fwd	GAAAAAAAATCTTAGGTGC		
Sall1-E3-Rev	GAGCAAACAACAGCCTTCCC		
Sall1-E4-Fwd	GTTTGTTCAATTTTTAAATT		
Sall1-E4-Rev	ACATTGGCCTAGAAGGTATC		
Sall1-E5-Fwd	CAGGGGAAGGAAGGCAGGCT		
Sall1-E5-Rev	GTGGGACCCTTGCCGGTGGC		

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

Genotyping Sall1-BAC insertion into ES cells			
Sex-PCR-Fwd	CTGAAGCTTTTGGCTTTGAG		
Sex-PCR-Rev	CCACTGCCAAATTCTTTGG		
5'ITR-BAC-Fwd	gacgcatgcattcttgaaat		
5'ITR-BAC-Rev	atgcgtcattttgactcacg		
3'ITR-BAC-Fwd	gaagaaattttgagtttttgttttt		
3'ITR-BAC-Rev	cgcatgtgttttatcggtct		
bck-BAC-Fwd	GGCGGTGTTGATACAGCGGGTAA		
bck-BAC-Rev	CCGGCGTTCGGTCGAAGAGTATC		