Supplemental Materials:

Materials and Methods

Mouse genetics

P0-3.9-GFPCre mice were described previously and maintained as heterozygotes on an FVB background (Rowan et al. 2008). *Pax6*$^{Sey\text{-}Neu}$ mice were maintained as heterozygotes on a C3H background. *Prep1*^{i} (gift of Dr. Francesco Blasi) and *Prep1*$^{-}$ (gift of Dr. Neal Copeland) mice were maintained as heterozygotes on a C57BL/6 background (Ferretti et al. 2006). Transgenic constructs were based on the pLNGLKS reporter plasmid containing the 526 bp Pax6 EE directing lacZ expression (Zhang et al. 2002), except for the P0-3.9-L1ΔL2-lacZ reporter, which was based on the P0-3.9-lacZ reporter line (Zhang et al. 2006). Site-directed mutations were constructed using the Quikchange® Lightning Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's directions. Linearized transgenic inserts were injected into the male pronuclei of fertilized FVB eggs using standard techniques. Staged embryos were collected and stained for β-galactosidase activity (Whiting et al. 1991). Images of whole-mount embryos were captured with a Leica dissecting microscope and DC500 digital camera or a Nikon inverted microscope for fluorescence imaging. All mouse work was performed in accordance with protocols approved by the Harvard Animal Care and Usage Committee.

Tissue Analyses

Embryonic tissue was fixed in 4% paraformaldehyde/PBS for 15 minutes at 4^{O}C and processed by stepwise sucrose/PBS incubation for 10 μm frozen sections in OCT by standard techniques. For antibody staining, primary antibodies used were mouse anti-Pax6 (1:10; Development Studies Hybridoma Bank, The University of Iowa), rabbit anti-Pax6 (1:1000; Covance Research Products), rabbit anti-Foxe3 (1:1000; gift of Dr. Peter Carlsson), goat anti-gamma-crystallin (1:1000; Santa Cruz Biotechnology), goat anti-Pdx1 (1:5000; gift of Dr. Chris Wright), guinea pig anti-Prep1 (1:200; (Zhang et al. 2006)), guinea pig anti-Six3 (1:500; Abcam), rabbit anti-Sox2 (1:1000; Chemicon), goat anti-Meis1/2 (1:250; Santa Cruz Biotechnology). Secondary antibodies were generated in donkey versus the appropriate species and directly conjugated with Cy3 (Jackson Immunologicals) or Alexa Fluor 488 (Molecular Probes). Nuclei were stained with DAPI (Sigma). Labeled sections were visualized with a Zeiss fluorescent microscope equipped with a Leica DFC350 digital camera. Standard histologic staining of frozen sections was also performed. All images were processed using Adobe Photoshop software (CS3) and manipulated electronically to adjust brightness and contrast as well as pseudocoloring.

Protein Expression and Surface Plasmon Resonance

GST-Prep1 and GST-Meis1 homeodomain open reading frames (Berger et al. 2008) or a full-length GST-Prep1 construct were expressed in *E. coli* (induction with 0.1 mM IPTG, followed by expression at 37°C for 2 hrs). Cell pellets were resuspended in lysis buffer (150 mM Tris-HCl, pH 8.0; 150 mM NaCl; 2 mM DTT; EDTA-free protease inhibitor; 1

mg/ml lysozyme) and lysed via freeze-thaw (3 cycles of 15 min dry ice / 15 min room temperature water). Filtered lysate was applied to GSTrap™ FF affinity columns (GE Healthcare) using an AKTA prime plus FPLC (GE Healthcare) and eluted off the column with elution buffer (10 mM reduced L-glutathione; 50 mM Tris-HCl, pH 8.0; 2 mM DTT; 0.02% Triton-X-100). Concentration of purified protein was determined by Coomassie Bradford assay.

Surface plasmon resonance (SPR) was performed on a Biacore 3000. Single-stranded DNA oligonucleotides (Integrated DNA Technologies) were double-stranded using a biotin-conjugated primer, and purified using MinElute spin columns (Qiagen). Biotinylated oligos were immobilized onto a Sensor Chip SA (Biacore); a target response of 30 RU was used during DNA immobilization step. Serial concentrations of protein sample were diluted in running buffer (10 mM Tris-HCl, pH 7.4; 3 mM DTT; 0.02% Triton-X-100; 0.2 mM EDTA; 120 mM NaCl) and applied to the Sensor Chip @ 25 uL/min, using the KINJECT option, 250 uL sample (1000 sec dissociation phase). Sensor Chip surface regeneration steps: 6 uL (1M NaCl); 6 uL (0.1 % SDS); 300 uL running buffer. Protein samples, at multiple concentrations, were applied to the chip two times and each was independently fit to both kinetic and saturation binding models using Scrubber2 software (BioLogic Software). Reported Kd values are the mean and standard deviation of these measurements. SPR probes sequences (L1, L2 and mutated sites are shown in bold; common primer sequence for L1, L2 and P1 probes is underlined):

L1: 5'-GTAGATCGAAGCCGGCC**TTGTCA**GGTTGAGAAAAAGTGAAG<u>AAAGGATGGGTGCGACGCG-3'</u>;

L2: 5'-TAACATCCAGGACGTGC**CTGTCT**ACTTTCAGAGAATTGCAGAAAGGATGGGTGCGACGCG-3';

P1: 5'-CCTAACGAGCCCTTTA**TTGACA**GACAGATAGATAAGCTGGGAAAGGATGGGTGCGACGCG-3';

L1L2: 5'-CGGCC**TTGTCA**GGTTGAGAAAAAGTGAATCTCTAACATCCAGGACGTGC**CTGTCT**ACTTG-3';

ΔL1L2: 5'-CGGCCAGACTAGGTTGAGAAAAAGTGAATCTCTAACATCCAGGACGTGC**CTGTCT**ACTTG-3'

Computational Analysis

The mouse genome sequence for the P0-3.9 region was downloaded from the UCSC genome repository (build mm8). The region was searched for Prep1 binding site sequences using custom Perl scripts (available by request); 8-mers were scored according to their PBM enrichment score, a statistical measure that ranges from -0.5 to +0.5 and indicates the binding preference of the protein, assayed in a universal PBM experiment, for a particular 8-mer as compared to all other 8-mers (Berger et al. 2008). We considered only those sequence matches for which the 8-mer PBM enrichment score was at least 0.37 (false discovery rate 0.001 (Badis et al. 2009)). Prep1 PBM data are available via the UniPROBE database (Newburger and Bulyk 2009). Search results were converted into GFF format and uploaded to the UCSC Genome Browser for visualization. All mutations generated in this study were tested for creation or deletion of known lens regulatory TF binding sites (using data from the UniPROBE database, at PBM enrichment score $\geq$ 0.37) and none were found.

Relative affinities from PBM data were calculated as the ratios of 8-mer median signal intensity values (Berger et al. 2008) (L1: CTTGTCAG; L2: CCTGTCTA; P1: TTGACAGA).

Enhancer Modeling

An equilibrium thermodynamic model, recently described by Bintu *et al.* (Bintu et al. 2005), was used to model transcriptional activation of the EE. Transcriptional activators bound at one or two sites (Supplemental Fig. 6A) recruit a transcriptional coactivator required for activation. In Bintu *et al.*, activators recruit the Pol II enzyme itself, but as we do not know what the Prep1 activator interacts with to facilitate transcription, we relax the definition to assume that Prep1 interacts with some required transcriptional cofactor.

While this model is perhaps an idealized representation of the complex biophysical interactions occurring at many mammalian promoters, we believe it may capture the essence of a DNA-bound transcription factor recruiting and/or stabilizing an integral component of the transcriptional machinery and thereby provide a useful quantitative description of transcriptional activation.

Model

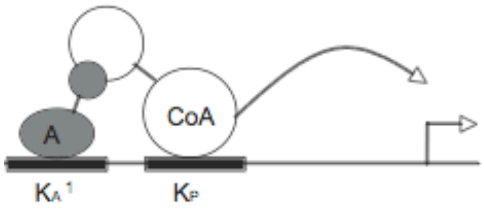
Transcriptional activation is modeled directly as the probability of the cofactor being bound to the promoter. The role of transcription factors is to modulate the binding of the cofactor, either positively or negatively, by direct interaction with it. The probability of the cofactor being bound (or alternatively interpreted as the fraction of maximal activation) is

$$P = \frac{1}{1+F} \qquad [1]$$

Relations for F that describe transcriptional activation via binding of a single species of transcription factor to one or two sites are shown below.

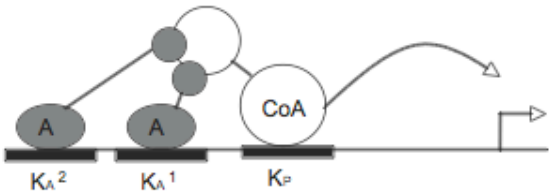
One binding site

$$F = \frac{\left(1+\frac{[A]}{K_A^1}\right)}{\left(1+\frac{[A]}{K_A^1 K_{AP}^1}\right)\frac{[P]}{K_P}} \qquad [2]$$



Two binding sites

$$F = \frac{\left(1 + \frac{[A]}{K_A^1} + \frac{[A]}{K_A^2} + \frac{[A][A]}{K_A^1 K_A^2 K_{AA}}\right)}{\left(1 + \frac{[A]}{K_A^1 K_{AP}^1} + \frac{[A]}{K_A^2 K_{AP}^2} + \frac{[A][A]}{K_A^1 K_A^2 K_{AP}^1 K_{AP}^2 K_{AA}}\right)\frac{[P]}{K_P}} \qquad [3]$$



For two binding sites with no synergy, set* $\mathbf{K_{AP}^1}$ $\mathbf{K_{AP}^2}$=**$\mathbf{K_{AP}}$

[A]: Nuclear concentration of Prep1 protein (*i.e.,* the activator *A*).

$\mathbf{K_A^i}$: Dissociation binding constant for Prep1 binding to site *i*. SPR-determined values of 92.3 nM and 106 nM were used for the affinities of the native sites L1 and L2, 14 nM was used for the high-affinity mutations (Fig. 3).

$\mathbf{K_{AA}}$: Dissociation binding constant describing the interaction between Prep1 molecules bound to sites L1 and L2. This quantifies the strength of any Prep1-Prep1 interaction leading to cooperative binding. As we do not observe cooperativity between Prep1 molecules in our Biacore experiments (Supplemental Fig. 5), we set this parameter value to 1 (*i.e.,* no cooperativity).

$\mathbf{K_{AP}^i}$: Dissociation binding constant describing the binding interaction *between* the Prep1 bound at site *i* and the cofactor. This quantifies the strength of cofactor recruitment by DNA-bound Prep1.

[P]: Nuclear concentration of required cofactor.

K_P: Dissociation binding constant of the cofactor binding to the promoter.

Three parameters remain undetermined: [**A],** the concentration of the Prep1 activator; **K_{AP}**, the recruitment strength; and the ratio **[P]/ K_P** which essentially defines the basal transcription level when no activator is present (i.e. setting [A] equal to 0, we see that P = $1/(1 + K_P/[P])$). We were unable to directly establish estimates for these remaining parameters; indeed, we do not know what factors Prep1 recruits in this developmental context, nor what the basal level of transcription is (*i.e.,* whether **K_P/[P] $_{should}$** be 4,000 or 10,000). Instead, we have utilized our reporter construct expression data (Figs. 3C, 4A, 4C) and examined 1) whether simple biophysical models based on formulae 2 and 3 would yield predictions consistent with these reporter data, and 2) what range of parameter values would provide consistent models. We view this as using our *in vivo* reporter constructs to indirectly parameterize our model.

Comparing Model Predictions to Reporter Expression Data at E10.5

Initially we examined reporter expression data for E10.5 embryos (Figs. 3C, 4A). We required that our model predictions met the following conservative criteria:

i) Expression levels of the native-site reporter construct (L1L2) should be at least 4-fold higher than levels for the high-affinity site mutant reporters (ΔL1L2*, L1*ΔL2);

ii) Expression levels of the high-affinity site mutant reporter constructs (ΔL1L2*, L1*ΔL2) should be at least 2-fold higher than the single-site mutant reporter constructs (ΔL1L2, L1ΔL2) (see also Fig. 4C solid grey line);

iii) Expression levels of the two high-affinity reporter constructs (L1*L2*) should be no more than ~30% (1.3-fold) higher than the native reporter construct (L1L2).

In Supplemental Figure 6, for a range of $\mathbf{K_{AP}}$ and $\mathbf{K_P}$ **/[P]** parameter values, we show modeled absolute (Supplemental Fig. 6A) and relative (Supplemental Fig. 6B) expression levels for our reporter constructs as a function of nuclear Prep1 protein concentration. Relative reporter expression levels are shown to more readily facilitate the comparison of our above reporter criteria with the model prediction. The orange shaded bands in Supplemental Figure 6 indicate the Prep1 concentration range over which the models are consistent with our reporter construct data for the indicated values of $\mathbf{K_{AP}}$ and $\mathbf{K_P}$ **/[P]**; the solid black box indicates those models for which such a solution consistent with the reporter expression data could be identified. Consistent models could not be identified for some parameter values, for example when the recruitment interaction energy ($\mathbf{K_{AP}}$) was too weak (top row, $\mathbf{K_{AP}}$ = 10 mM).

Without additional data we cannot formally decide between these parameter values; however, published data for other systems suggest further refinement of the parameter

values. Buchler *et al.* (Buchler et al. 2003) describe the weak interactions used by transcription factors to recruit cofactors and transcription machinery as a range from ~1 to 4 kcal/mol. This translates to $\mathbf{K_{AP}}$ values of ~180 mM to 1.2 mM, respectively, in our model. As illustrated in Supplemental Figure 6, we were unable to find consistent models when the values of $\mathbf{K_{AP}}$ become too high (*i.e.,* recruitment became too weak); no consistent models could be found for $\mathbf{K_{AP}}$ values of 10 mM or greater when $\mathbf{K_P/[P]}$ values were tested exhaustively at intervals of 2-fold from 10 to 100,000. This suggests a value of around 1 mM for $\mathbf{K_{AP}}$. Of note, for models with $\mathbf{K_{AP}}$ = 1 mM (second row of Supplemental Fig. 6) the predicted range of Prep1 protein concentration ranges from ~10 to 100 nM. This range of Prep1 concentration is broadly consistent with estimates of the nuclear concentration of Bicoid (also a homeodomain protein) by Gregor *et al.* (Gregor et al. 2007) in the developing fly embryo or NFκB in TNFα-responsive cell lines (Giorgetti et al. 2010). Other studies have suggested recruitment affinities ($\mathbf{K_{AP}}$) on the order of 10 to 100 μM (Farrell et al. 1996; Giangrande et al. 2000). These values yield models consistent with our *in vivo* data, provided that the $\mathbf{K_P/[P]}$ ratio increases concomitantly. These different modeling possibilities highlight a balance between basal transcriptional activation levels and activator recruitment strength. Based on the above refinements, parameter values of 1 mM and 4,000 were chosen for $\mathbf{K_{AP}}$ and $\mathbf{K_P/[P]}$, respectively, to generate the representative model shown in Figure 4.

Two features common to the class of consistent models provided insight into the EE/Prep1 *cis*-regulatory system. First, the predicted concentration range of Prep1

overlaps the steepest portion of the L1L2 expression curve, where it is highly sensitive to Prep1 concentration. This prediction is consistent with the sensitivity of *Pax6* expression observed in response to Prep1 dosage mutants (see main text). Second, the largest difference in expression between the L1L2 and L1*L2* constructs is predicted to occur below the predicted Prep1 concentration range. This suggested to us the hypothesis that the lower-affinity Prep1 sites have been conserved in order to prevent expression when Prep1 levels are lower, earlier in development (see main text).

Comparing Model Predictions to Additional Reporter Expression Data at E9.5

Reporter construct expression data at E9.5 for L1L2 and L1*L2* (Fig. 4D) provided additional constraints on our model and implicitly yielded a prediction about the Prep1 concentration at E9.5. These additional constraints were:

i) L1L2 levels should be, conservatively, at least 8-fold less at E9.5 than predicted for E10.5 (we use the lowest predicted Prep1 levels for E10.5);

ii) L1*L2* expression levels at E9.5 cannot be less than 50% of their E10.5 levels.

Supplemental Figure 7 shows the same model as in Figure 4C ($\mathbf{K_{AP}}$ = 1 mM and $\mathbf{K_P/[P]}$ = 4,000) with a blue bar delineating the predicted concentration range of nuclear Prep1

protein in the low nanomolar range that is consistent with our E9.5 reporter construct constraints.

Supplemental Table and Figure Legends

Supplemental Table 1: *Prep1* mutant mice show ocular phenotypes. The *Prep1*$^{-}$ allele refers to a targeted deletion of the gene, while the *Prep1*^{i} hypomorphic allele refers to a retroviral insertion in the first intron of *Prep1* leading to a 98% reduction in RNA expression. *Prep1*$^{+/-}$ heterozygotes were crossed to *Prep1*$^{+/i}$ heterozygotes to generate four different genotypes. Embryos were harvested between E10.5 and E12.5 for phenotypic analysis. Affected animals minimally had ocular phenotypes (small eyes or absent eyes), which frequently segregated with severe neural, craniofacial, and hematopoietic phenotypes. For *Prep1*$^{i/-}$ trans-heterozygotes, 17/75 embryos showed no apparent abnormalities. Eye phenotypes were quantified on a per-eye basis. In some cases, there were unilateral eye phenotypes with an affected embryo showing both a normal and small or absent eye. By reducing the genetic dosage of *Prep1*$^{i/i}$ by half, it is striking that we generate a more severe and penetrant phenotype, sometimes resulting in failure to gastrulate, as observed in Prep1$^{-/-}$ embryos (Ferretti et al. 2006). It is just as striking that expressing approximately 1% of the normal Prep1 RNA levels (based on (Ferretti et al. 2006)) can rescue embryonic viability from approximately E7.5 in *Prep1*$^{-/-}$ homozygotes to E12.5-E14.5 in *Prep1*$^{i/-}$ transheterozygotes, with a minority of embryos showing no phenotypic abnormalities at all. This indicates dramatic gene dosage sensitivity for *Prep1*.

Supplemental Figure 1: Prep1 is co-expressed with Pax6 in the presumptive lens ectoderm. Sections through E9.5 control or *Prep1*$^{i/-}$ mutants were stained with antibodies to Prep1 and Pax6. Prep1 immunostaining was ubiquitous in the eye region and largely, though not exclusively, nuclear. In this experiment, the chosen *Prep1*$^{i/-}$ mutant was phenotypically unaffected, although the protein is largely undetectable. Cells staining for both Prep1 and Pax6 appear as yellow or greenish-orange.

Supplemental Figure 2: Prep1 expression does not depend on Pax6 function. Sections through 16-somite E9.25 control or *Pax6*$^{Sey\text{-}Neu/Sey\text{-}Neu}$ mutants were stained with antibodies to Prep1 and Pax6. Prep1 immunostaining was both nuclear and cytoplasmic.

Supplemental Figure 3: Expression of known *Pax6* EE regulators are maintained in *Prep1* mutants. Sections through E10 control or *Prep1*$^{i/-}$ mutants were stained with antibodies to the transcription factors Six3, Sox2, or Meis1/2. *Prep1*$^{i/-}$ mutants did not show a thickened lens placode, but still retained expression of the above transcription factors in the presumptive lens ectoderm. Therefore, loss of Pax6 and Foxe3 expression was not caused by loss of expression of Six3, Sox2, or Meis1/2.

Supplemental Figure 4: Subsets of Prep1 binding sites are conserved across multiple vertebrates. Shown in the UCSC Genome Browser view are protein-binding microarray scores from a sliding 8-bp window set (PBM enrichment score ≥ 0.37). Conserved PBM hits are labeled with arrows. Hits in the minimal essential region of the EE (cyan) or

Pancreas enhancer (magenta) are labeled L1, L2, and P1 and the sequences are shown boxed below. Red nucleotides indicate divergence from the boxed CTGTCA core sequence. The full intervening sequences between L1 and L2, as well as flanking sequences around L1, L2, and P1, are shown.

Supplemental Figure 5: SPR measurements indicate that full-length Prep1 does not bind cooperatively to the L1L2 region. **(A)** SPR dose response curves for GST-full-length-Prep1 binding to ΔL1L2 (cyan, diamonds) compared to L1L2 (magenta, squares). **(B)** Kd values (mean and standard deviation) derived from the duplicate SPR concentration series shown in (A). The cooperative binding curve would be expected to have a steeper profile than a non-cooperative binding curve and a sharper transition from bound to unbound over a comparable concentration range. Since we do not see a change in the profile of the curve upon mutation of the L1 binding site, this demonstrates lack of cooperative binding to the L1L2 binding sites.

Supplemental Figure 6: Modeled reporter expression levels for a range of $\mathbf{K_{AP}}$ and $\mathbf{K_P/[P]}$ parameter values. $\mathbf{K_{AP}}$ is the dissociation binding constant describing the binding interaction between the Prep1 bound at a given site and the cofactor; *i.e.,* this quantifies the strength of cofactor recruitment by DNA-bound Prep1. $\mathbf{K_P}$ is the dissociation binding constant of the cofactor binding to the promoter. **[P]** is the nuclear concentration of required cofactor. Models in different columns refer to different set values for $\mathbf{K_P/[P]}$ ranging from 1,000 to 64,000 while models in different rows refer to different set values

for $\mathbf{K_{AP}}$ ranging from 10 μM to 10 mM. Expression levels are shown as a function of nuclear Prep1 protein concentration. Prep1 concentration ranges for which the model predictions agree with our reporter criteria (see text) are highlighted in orange. (**A)** Absolute expression levels for four different reporter constructs. (**B**) Relative expression levels for three different pairs of reporter constructs shown in (A). Model parameterization yielding predictions consistent with our reporter data are indicated with a solid black box. The parameters chosen for our representative model are indicated with a dashed black box.

Supplemental Figure 7: Modeled reporter expression levels highlighting predicted Prep1 concentration ranges for E10.5 and E9.5 embryos. Values of 1 mM and 4,000 were used for $\mathbf{K_{AP}}$ and $\mathbf{K_P/[P]}$, respectively. Absolute and relative expression levels for different reporter constructs are shown in the upper and lower panels, respectively. Predicted Prep1 concentration ranges over which the modeled expression levels are consistent with our reporter construct data for both E10.5 and E9.5 embryos are shaded in orange (E10.5) and blue (E9.5).

Supplemental Figure 8: Ectopic staining in the brain (arrowheads) and expanded surface ectoderm staining (arrows) of L1*L2*transgenic embryos ranging from E9.25 – E10.5 compared to control E9.5 L1L2 transgenic embryo.

References for Supplemental Materials

Badis, G., Berger, M.F., Philippakis, A.A., Talukder, S., Gehrke, A.R., Jaeger, S.A., Chan, E.T., Metzler, G., Vedenko, A., Chen, X. et al. 2009. Diversity and complexity in DNA recognition by transcription factors. *Science* **324**(5935): 1720-1723.

Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Pena-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T. et al. 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**(7): 1266-1276.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. 2005. Transcriptional regulation by the numbers: models. *Curr Opin Genet Dev* **15**(2): 116-124.

Buchler, N.E., Gerland, U., and Hwa, T. 2003. On schemes of combinatorial transcription logic. *Proc Natl Acad Sci USA* **100**(9): 5136-5141.

Farrell, S., Simkovich, N., Wu, Y., Barberis, A., and Ptashne, M. 1996. Gene activation by recruitment of the RNA polymerase II holoenzyme. *Genes Dev* **10**(18): 2359-2367.

Ferretti, E., Villaescusa, J.C., Di Rosa, P., Fernandez-Diaz, L.C., Longobardi, E., Mazzieri, R., Miccio, A., Micali, N., Selleri, L., Ferrari, G. et al. 2006. Hypomorphic mutation of the TALE gene Prep1 (pKnox1) causes a major

reduction of Pbx and Meis proteins and a pleiotropic embryonic phenotype. *Mol Cell Biol* **26**(15): 5650-5662.

Giangrande, P.H., Kimbrel, E.A., Edwards, D.P., and McDonnell, D.P. 2000. The opposing transcriptional activities of the two isoforms of the human progesterone receptor are due to differential cofactor binding. *Mol Cell Biol* **20**(9): 3102-3115.

Giorgetti, L., Siggers, T., Tiana, G., Caprara, G., Notarbartolo, S., Corona, T., Pasparakis, M., Milani, P., Bulyk, M.L., and Natoli, G. 2010. Noncooperative interactions between transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs. *Mol Cell* **37**(3): 418-428.

Gregor, T., Tank, D.W., Wieschaus, E.F., and Bialek, W. 2007. Probing the limits to positional information. *Cell* **130**(1): 153-164.

Newburger, D. and Bulyk, M.L. 2009. UniPROBE: an online database of protein binding microarray data on protein–DNA interactions. *Nucleic Acids Res* **37**: D77-82.

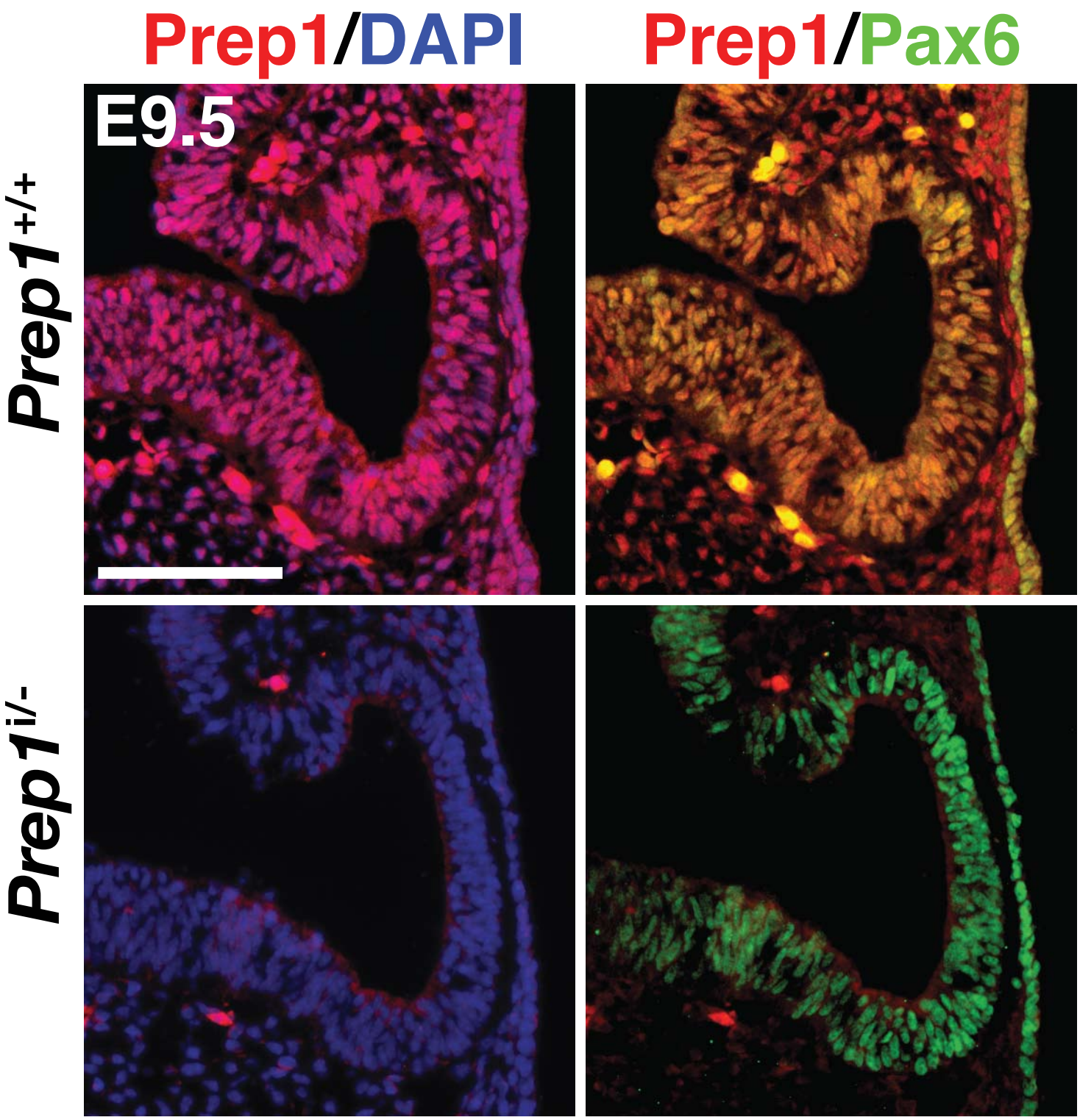
Rowan, S., Conley, K.W., Le, T.T., Donner, A.L., Maas, R.L., and Brown, N.L. 2008. Notch signaling regulates growth and differentiation in the mammalian lens. *Dev Biol* **321**(1): 111-122.

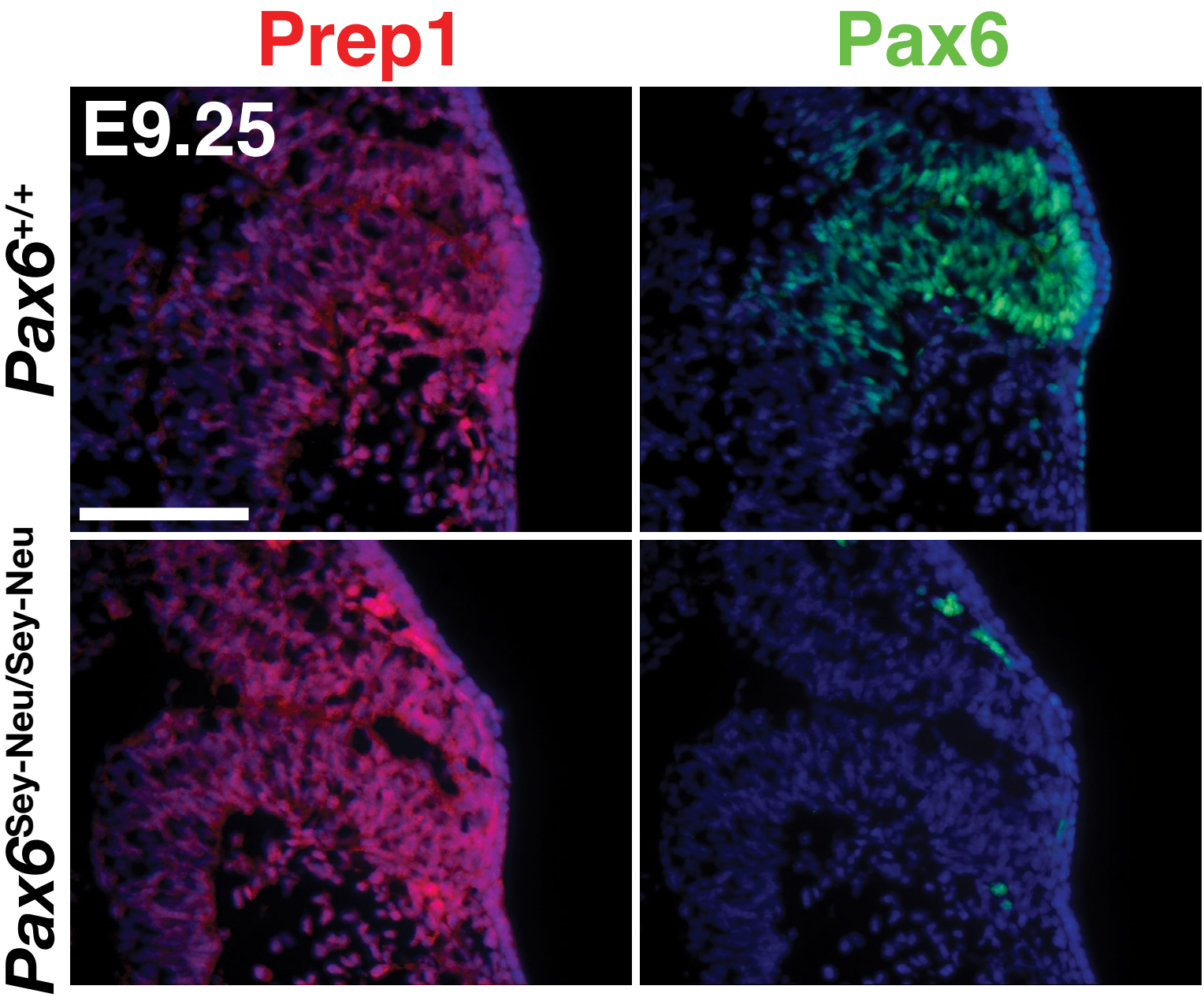
Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P.W., Stott, D., and Allemann, R.K. 1991. Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. *Genes Dev* **5**(11): 2048-2059.

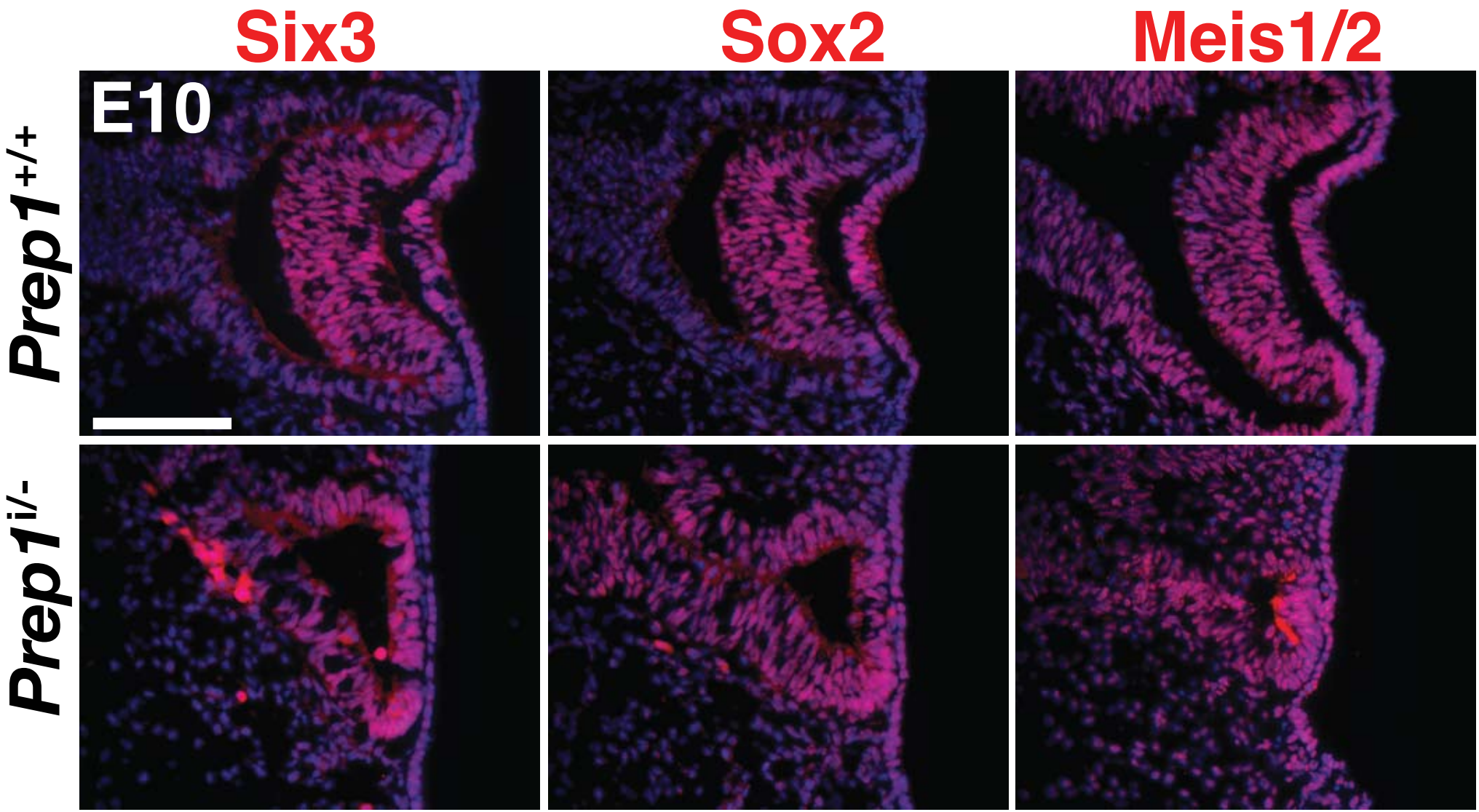
Zhang, X., Friedman, A., Heaney, S., Purcell, P., and Maas, R.L. 2002. Meis homeoproteins directly regulate Pax6 during vertebrate lens morphogenesis. *Genes Dev* **16**(16): 2097-2107.

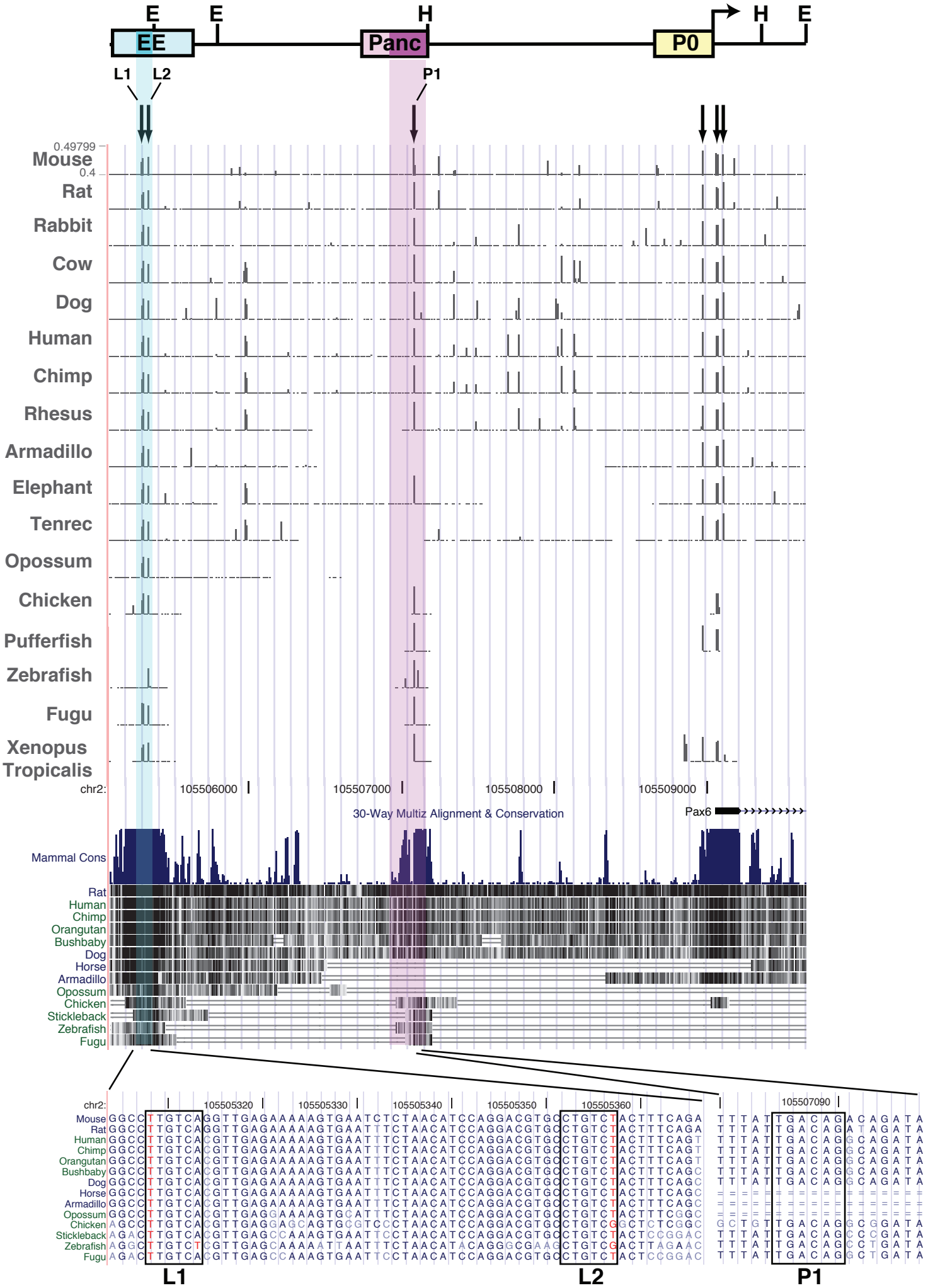
Zhang, X., Rowan, S., Yue, Y., Heaney, S., Pan, Y., Brendolan, A., Selleri, L., and Maas, R.L. 2006. Pax6 is regulated by Meis and Pbx homeoproteins during pancreatic development. *Dev Biol* **300**(2): 748-757.

			Eye phenotypes:		
Genotype	Number of animals	Affected animals	Normal	Small	Absent
Prep1$^{+/+}$	82	1	163	0	1
Prep1$^{+/-}$	89	2	175	1	2
Prep1$^{+/i}$	76	2	148	1	3
Prep1$^{i/-}$	75	58	41	27	82

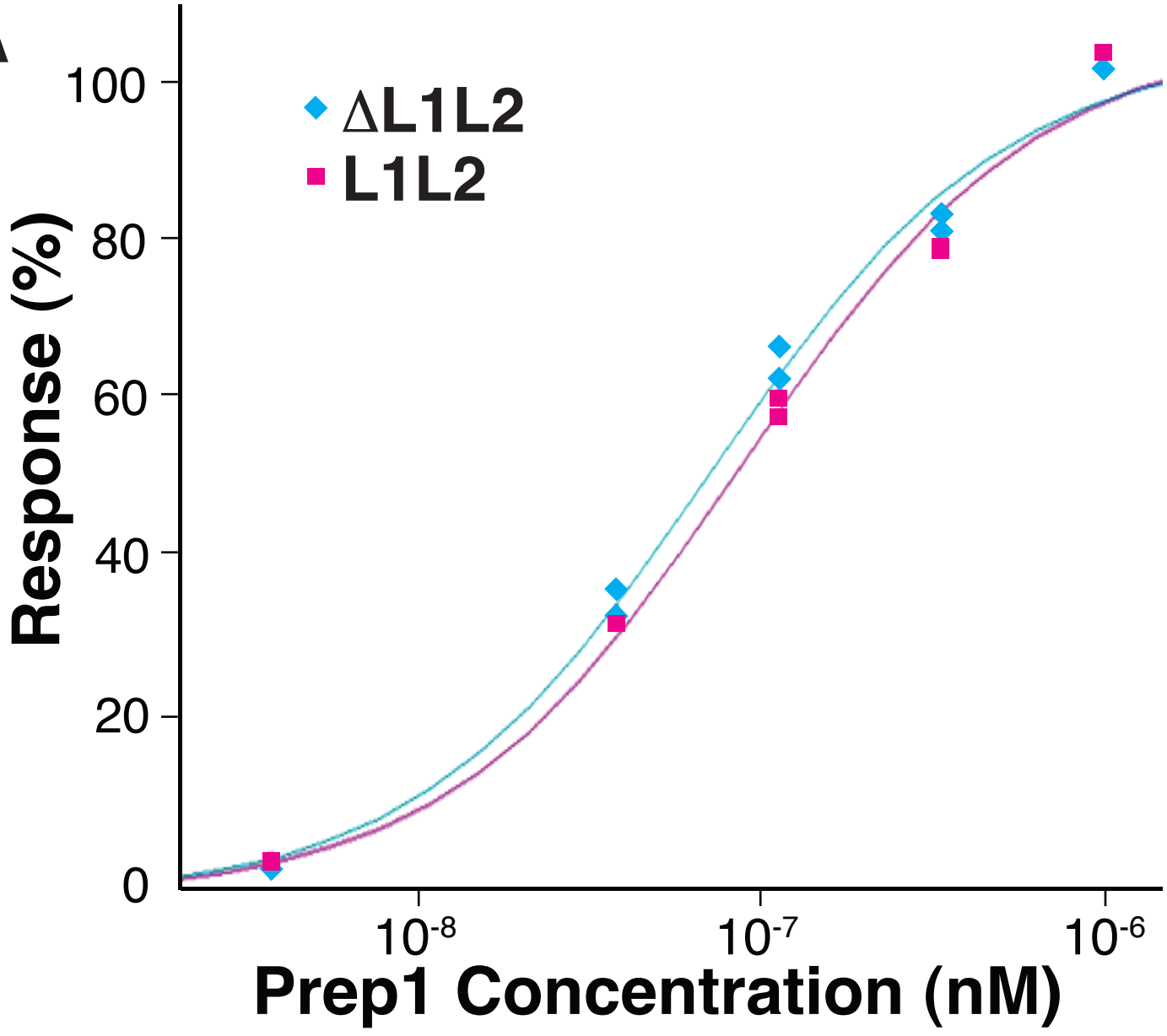

Prep1/DAPI
Prep1/Pax6
E9.5
Prep1+/+
Prep1i/-


Prep1
Pax6
E9.25
Pax6+/+
Pax6Sey-Neu/Sey-Neu


Six3
Sox2
Meis1/2
E10
Prep1+/+
Prep1i/-

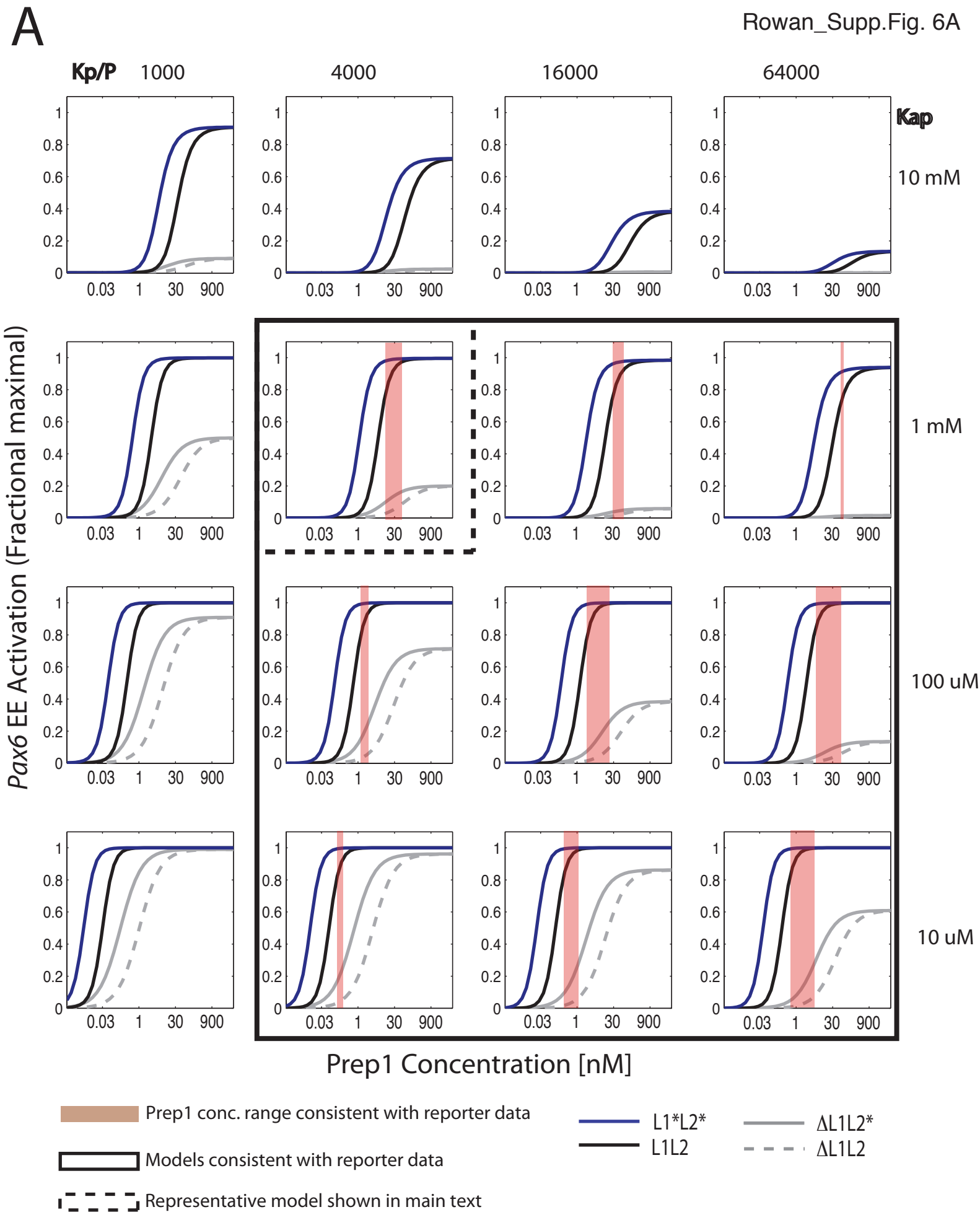

E
E
H
H
E
EE
Panc
P0
L1
L2
P1
0.49799
0.4
Mouse
Rat
Rabbit
Cow
Dog
Human
Chimp
Rhesus
Armadillo
Elephant
Tenrec
Opossum
Chicken
Pufferfish
Zebrafish
Fugu
Xenopus Tropicalis
chr2:
105506000
105507000
105508000
105509000
30-Way Multiz Alignment & Conservation
Pax6
Mammal Cons
Rat
Human
Chimp
Orangutan
Bushbaby
Dog
Horse
Armadillo
Opossum
Chicken
Stickleback
Zebrafish
Fugu
chr2:
105505320
105505330
105505340
105505350
105505360
105507090
Mouse GGCCTTGTCAGGTTGAGAAAAAGTGAATCTCTAACATCCAGGACGTGCCTGTCTACTTTCAGA TTTATTGACAGACAGATA
Rat GGCCTTGTCAGGTTGAGAAAAAGTGAATTTCTAACATCCAGGACGTGCCTGTCTACTTTCAGA TTTATTGACAGATAGATA
Human GGCCTTGTCACGTTGAGAAAAAGTGAATTTCTAACATCCAGGACGTGCCTGTCTACTTTCAGT TTTATTGACAGGCAGATA
Chimp GGCCTTGTCACGTTGAGAAAAAGTGAATTTCTAACATCCAGGACGTGCCTGTCTACTTTCAGT TTTATTGACAGGCAGATA
Orangutan GGCCTTGTCACGTTGAGAAAAAGTGAATTTCTAACATCCAGGACGTGCCTGTCTACTTTCAGT TTTATTGACAGGCAGATA
Bushbaby GGCCTTGTCACGTTGAGAAAAAGTGAATTTCTAACATCCAGGACGTGCCTGTCTACTTTCAGC TTTATTGACAGGCAGATA
Dog GGCCTTGTCACGTTGAGAAAAAGTGAATTTCTAACATCCAGGACGTGCCTGTCTACTTTCAGC TTTATTGACAGGCAGATA
Horse GGCCTTGTCACGTTGAGAAAAAGTGAATTTCTAACATCCAGGACGTGCCTGTCTACTTTCAGC = = = = = = = = = = = = = = =
Armadillo GGCCTTGTCACGTTGAGAAAAAGTGAATTTCTAACATCCAGGACGTGCCTGTCTACTTTCAGC = = = = = = = = = = = = = = =
Opossum GGCCTTGTCACGTTGAGGAAAAAGTGCATTTCTAACATCCAGGACGTGCCTGTCTACTTTCGGC = = = = = = = = = = = = = = =
Chicken AGCCTTGTCACGTTGAGGAGCAGTGCGTCCCTAACATCCAGGACGTGCCTGTCGGCTCTCGGC GCTGTTGACAGGCGGATA
Stickleback AGACTTGTCACGTTGAGCCAAAGTGAATTCCTAACATCCAGGACGTGCCTGTCTACTCCGGAC TTTATTGACAGCCCGATA
Zebrafish AGGCTTGTCTCGTTGAGCAAAAATTAATTTCTAACATACAGGGCGAAGCTGTCGACTTAGAAC TTTATTGACAGCCTGATA
Fugu AGACTTGTCACGTTGAGCCAAAGTGAATTCCTAACATCCAGGACGTGCCTGTCTACTCCGGAC TTTATTGACAGGCTGATA
L1
L2
P1

A



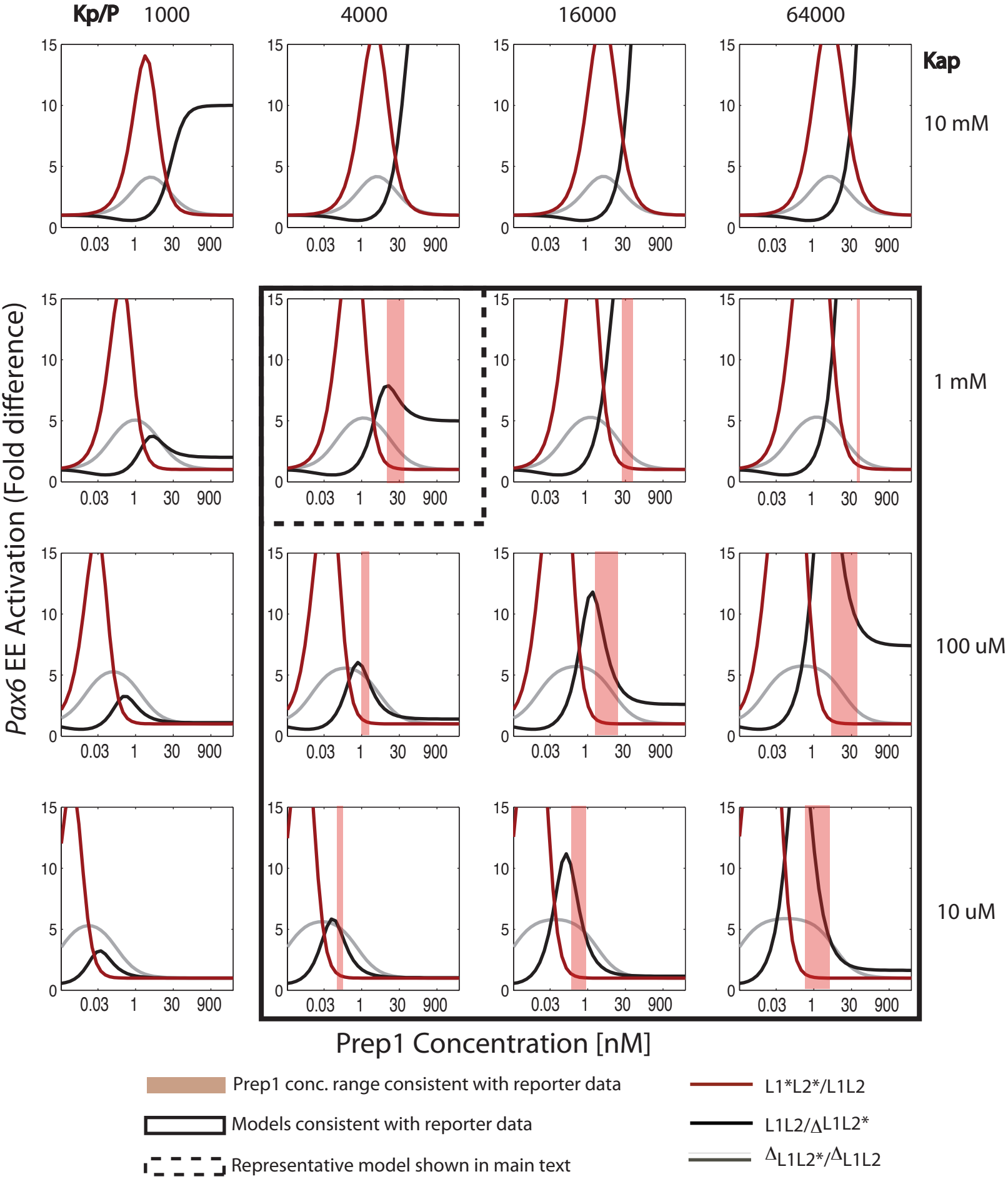
B

Construct	Kd (nM)	SD
ΔL1L2	54	6
L1L2	67	3

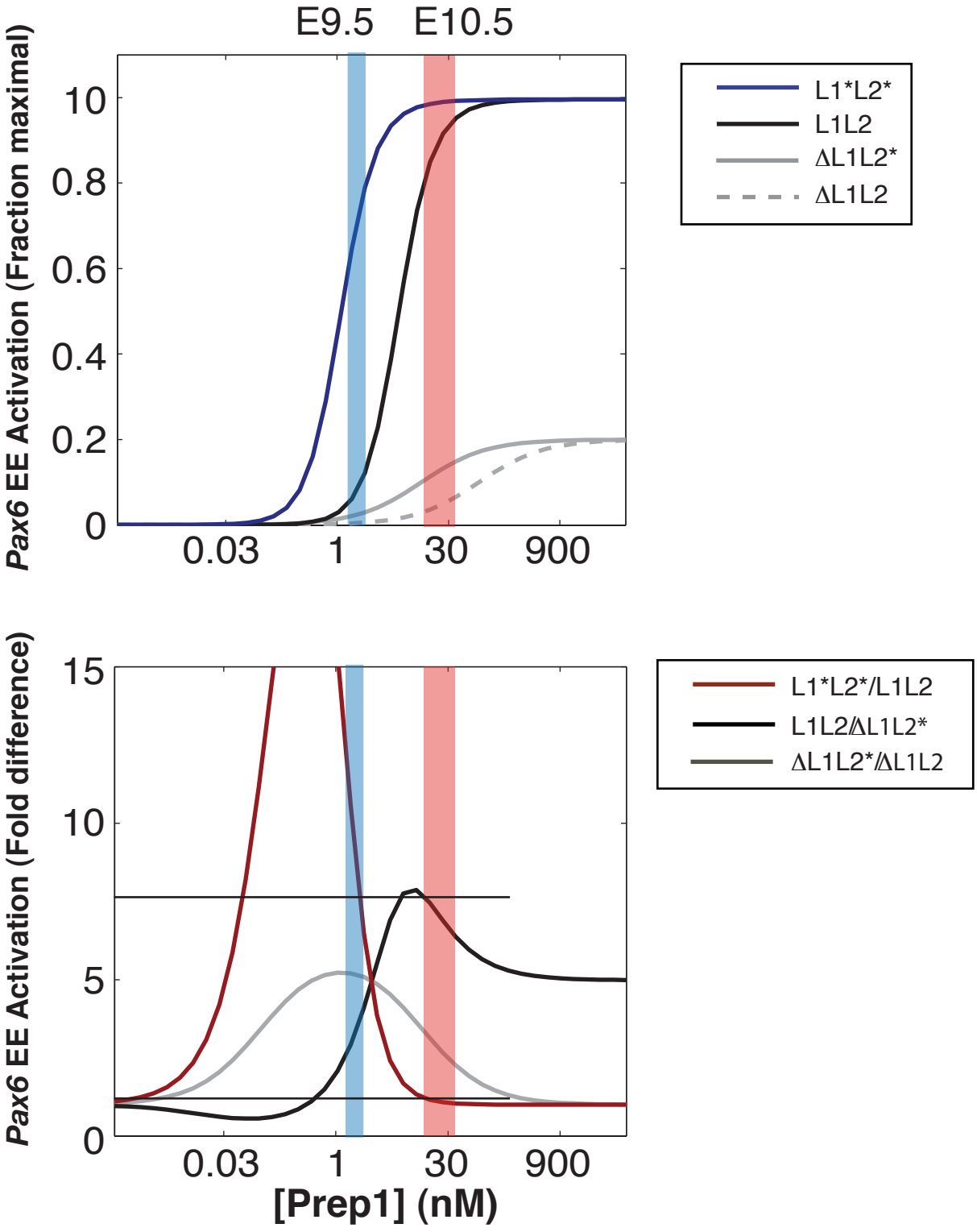

A
Rowan_Supp.Fig. 6A
Kp/P
1000
4000
16000
64000
Kap
10 mM
1 mM
100 uM
10 uM
Pax6 EE Activation (Fractional maximal)
Prep1 Concentration [nM]
0.03
1
30
900
Prep1 conc. range consistent with reporter data
Models consistent with reporter data
Representative model shown in main text
L1*L2*
L1L2
ΔL1L2*
ΔL1L2

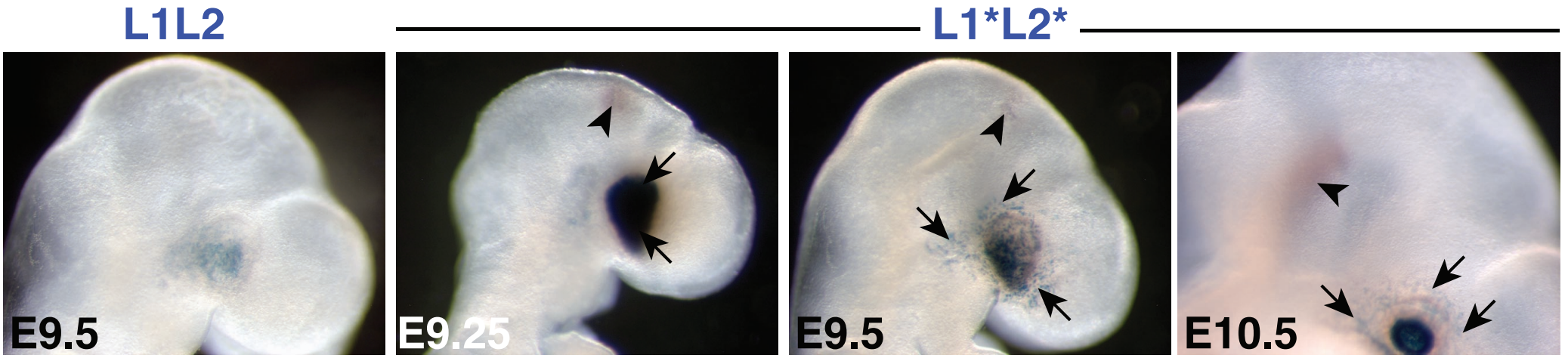
B

Rowan_Supp.Fig. 6B



Predicted Prep1 Concentration Ranges




L1L2
L1*L2*
E9.5
E9.25
E9.5
E10.5