Figure S1. Validated pax6 DNA binding sites display a high level of dissimilarity. Experimentally validated pax6 binding sites were manually annotated from the literature. These are highly variable as illustrated by their Clustal alignment. Publications where sequences were identified: (1) PMID 9224716;(2) PMID 10051657; (3) PMID 11222774;(4) PMID 12783797; (5) PMID 16023139; (6) PMID 10094478; (7) PMID 12951074; (8) PMID 7753863; (9) PMID 16023139; (10) PMID 8798491; (11) PMID 9710641; (12) PMID 14732405; (13) PMID 12710953; (14) PMID 11943482; (15) PMID 9710641; (16) PMID 8798491; (17) PMID 15180990; (18) PMID 15110720; (19) PMID 12167158; (20) PMID 15522290; (21) PMID 11222774; (22) PMID 9710641; (23) PMID 12783797; (24) PMID 15180990; (25) PMID 14673159; (26) PMID 11222774; (27) PMID 16115881; (28) PMID 15161828; (29) PMID 9224716

Figure S2. An iterative procedure was used to identify a core subset of similar pax6 binding sites.
(A) Using manual observation, an initial set of Pax6 BSs was extracted and used to generate pax6 HMMs that were used to characterize the remaining pax6 BSs. Manual curation of these results allowed the addition of more Pax6 BSs to the original set and the start of a new iteration of the procedure. The procedure stopped when no new "good" BSs were identified by the current version of HMMs. *(B)* The resulting set of 16 Pax6 BSs show a high degree of similarity, although not identity.

Figure S3. A Comparison between pax6HMM and pax6PWM approaches.
(A) The PWM and HMM (above threshold) results were split into score percentiles and these were used to show that almost all HMM^{abv} are in the top 10% of the PWM results. *(B)* Very few HMM^{abv} hits are not also PWM hits. *(C)* Reciprocally, the PWM hits that are not in the HMM^{abv} set are uniformly distributed according to PWM scores. *(D)* The same analysis using all HMM results shows that most of the best PWM hits are still in the HMM set, but not all score high for the HMM method.

Figure S4. Gene ontology over-representation for the *in silico* mouse predictions. The x-axis represents the number of genes for each of the most highly over-represented molecular function GO terms, which are named on the y-axis and the p-values shown. The set of target genes is enriched in genes with transcription factor activity.

Figure S5. *pax6(a+b)* morpholino oligonuceotides down-regulate pax6 protein levels *in vivo*.
Protein extracts were made from pax6(a+b)MO and control embryos, at 28 hpf and analyzed by western blot for pax6 levels. pax6 protein levels were reduced on morpholino treatment. gapdh was used as loading control

Figure S6. pax6a or/and pax6b knockdown disrupts the expression of putative target genes in the neural tube.
Of the fifteen putative targets analyzed by WMISH, three (*maf*, *gata3* and *pax6b*) displayed altered expression patterns at 28 hpf in the neural tube, as indicated by full black arrow heads. In particular expression of *pax6b* seems to be up regulated when pax6 is down regulated.

Figure S7. pax6a or/and pax6b knockdown disrupts the expression of putative target genes in the midbrain and hindbrain.
Of the fifteen putative targets analyzed by WMISH, two (*pax6b* and *ptf1a*) displayed altered expression patterns at 28 hpf within the midbrain and hindbrain, as indicated by full black arrow heads.

Figure S8. Analysis of the coverage of intergenic regions.
The distances between the analysed loci and their closest upstream (downstream) gene (total) or evolutionarily conserved gene (cons) were computed split into 10kb bins. 80% of all genes are within 40 kb of their nearest neighbour.

Table S1. Oligonucleotide sequences used to produce antisense RNA *in situ* probes for zebrafish.

Table S2. *pax6a* and *pax6b* morpholinos used for zebrafish knock-down

Table S3. qPCR primers used to assess ChIP target enrichment

Table S4. Putative numbers of target genes and binding sites (enhancers).
For each cross-species comparison (human-zebrafish or mouse-zebrafish) the numbers represent mammalian targets with zebrafish targets in parentheses.

Table S5. Non-overlapping human PAX6 ECRs and neighbouring genes.

Table S6. Non-overlapping mouse PAX6 ECRs.

Table S7. Non-overlapping human-coincident PAX6 ECRs in zebrafish.

Table S8. Non-overlapping mouse-coincident PAX6 ECRs in zebrafish.
(tables S5-S8 are in annexed files)

Table S9. Some of the predicted target genes associated with human disease or mouse model with relevant phenotype