Differential regulation of mouse and human nephron progenitors by the Six family of transcriptional regulators

Lori L. O’Brien¹, Qiuyu Guo¹,², YoungJin Lee¹,*, Tracy Tran¹, Jean-Denis Benazet¹,⁻, Peter H. Whitney¹, Anton Valouev²,§ and Andrew P. McMahon¹,§

ABSTRACT

Nephron endowment is determined by the self-renewal and induction of a nephron progenitor pool established at the onset of kidney development. In the mouse, the related transcriptional regulators Six1 and Six2 play non-overlapping roles in nephron progenitors. Transient Six1 activity prefigures, and is essential for, active nephrogenesis. By contrast, Six2 maintains later progenitor self-renewal from the onset of nephrogenesis. We compared the regulatory actions of Six2 in mouse and human nephron progenitors by chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq). Surprisingly, SIX1 was identified as a SIX2 target unique to the human nephron progenitors. Furthermore, RNA-seq and immunostaining revealed overlapping SIX1 and SIX2 activity in 16 week human fetal nephron progenitors. Comparative bioinformatic analysis of human SIX1 and SIX2 ChIP-seq showed each factor targeted a similar set of cis-regulatory modules binding an identical target recognition motif. In contrast to the mouse where Six2 binds its own enhancers but does not interact with DNA around SIX1, both human SIX1 and SIX2 bind homologous SIX1 enhancers and putative enhancers positioned around SIX1. Transgenic analysis of a putative human SIX1 enhancer in the mouse revealed a transient, mouse-like, pre-nephrogenic, SIX1 regulatory pattern. Together, these data demonstrate a divergence in SIX-factor regulation between mouse and human nephron progenitors. In the human, an autocross-regulatory loop drives continued SIX1 and SIX2 expression during active nephrogenesis. By contrast, the mouse establishes only an auto-regulatory SIX2 loop. These data suggest differential SIX-factor regulation might have contributed to species differences in nephron progenitor programs such as the duration of nephrogenesis and the final nephron count.

KEY WORDS: Nephrogenesis, Nephron, Regulatory network, SIX1/2, Transcription

INTRODUCTION

Nephrons are the major functional unit of the kidney, filtering the blood to eliminate waste products, maintaining water, salt and pH balance, and regulating blood volume and pressure. A typical human kidney is composed of approximately one million nephrons, although this number ranges considerably (Bertram et al., 2011).

The final human nephron number is established prior to birth; nephrogenesis is reported to cease around 36 weeks of gestation (Hinchcliffe et al., 1991). Altered renal function and reduced nephron numbers are associated with premature birth and low birth weight, respectively (Mañalich et al., 2000; Rodriguez-Soriano et al., 2005; Hughson et al., 2003). Several studies have shown a link between low nephron number and an increased risk of hypertension late in life (Brenner et al., 1988; Keller et al., 2003; Hughson et al., 2006). An understanding of the determinants of nephron number might facilitate prevention of kidney and kidney-related disease.

In the mouse, all nephrons are derived from a pool of self-renewing metanephric mesenchyme progenitors established around embryonic day (E)10–E10.5 (Kobayashi et al., 2008). This population surrounds the invading epithelial ureteric bud tips of the nascent collecting duct at E11.0 and commences nephrogenesis (Boyle et al., 2008; Kobayashi et al., 2008). At each round of ureteric branching, nephrons are induced by a Wnt9b signal emanating from the ureteric epithelium (Carroll et al., 2005). Wnt9b and other factors also promote the expansion of the progenitor pool (Self et al., 2006; Kobayashi et al., 2008; Karner et al., 2011; Barak et al., 2012; Xu et al., 2014), which undergoes a large increase over the course of nephrogenesis (Short et al., 2014). The nephron progenitor pool persists until postnatal day (P)2–P3; its depletion marks the cessation of nephrogenesis with the generation of around 13,000 nephrons over a 12 day period of active kidney development (Hartman et al., 2007; Rumballe et al., 2011; Cullen-McEwen et al., 2003). Several transcriptional regulators are crucial for establishing or maintaining this population, including Sall1, Wt1, Oya1, Pax2, Hox11 paralogs, and two closely related Six-family members, Six1 and Six2 (Kreidberg et al., 1993; Torres et al., 1995; Xu et al., 1999, 2003, 2014; Nishinakamura et al., 2001; Wellik et al., 2002; Li et al., 2003; James et al., 2006; Self et al., 2006; Xu and Xu, 2015).

The founding member of the Six family, sine oculus (so), was first discovered in Drosophila melanogaster, where analysis of mutants established so as a major regulator of visual system development (Milani, 1941; Fischbach and Heisenberg, 1981; Fischbach and Technau, 1984; Cheyette et al., 1994; Serikaku and O’Tousa, 1994). Subsequent studies identified two additional family members, optix (also known as D-Six3) and D-Six4, with roles in eye development and mesoderm derivatives, respectively (Seo et al., 1999; Seimiya and Gehring, 2000; Kirby et al., 2001; Kenyon et al., 2005; Clark et al., 2006; Weasner et al., 2007). Vertebrate homologs have been characterized for all three founding family members, optix (also known as D-Six3) and D-Six4, with roles in eye development and mesoderm derivatives, respectively (Seo et al., 1999; Seimiya and Gehring, 2000; Kirby et al., 2001; Kenyon et al., 2005; Clark et al., 2006; Weasner et al., 2007).

Vertebrate homologs have been characterized for all three founding members and reveal an additional duplication of each Six gene, giving rise to six mammalian members: Six1-Six6. On the basis of sequence analysis and gene structure, Six1 and Six2 diverged from so, Six3 and Six6 from optix, and Six4 and Six5 from D-Six4 (Seo et al., 1999). Six factors bind DNA through a conserved homeodomain whereas the shared Six domain facilitates
interactions with co-regulators such as eya/Eya1 (Pignoni et al., 1997; Seo et al., 1999). Despite the divergence of Six1 and Six2 from so, neither gene is expressed or functions in the developing mouse eye. Instead Six1 and Six2 are expressed in a number of other developing tissues including the otic placode, branchial arches, muscle and kidney (Oliver et al., 1995).

In the developing mouse kidney, transient Six1 activity in the early kidney rudiment at E10.5 is essential for ureteric bud outgrowth and metanephric mesenchyme survival (Xu et al., 2003; Li et al., 2003; Xu and Xu, 2015) whereas sustained Six2 activity in the nephron progenitors is essential for their self-renewal, acting, at least in part, to block progenitor commitment to nephrogenesis (Self et al., 2006; Kobayashi et al., 2008; Park et al., 2012). Consequently, a loss-of-function for either gene results in kidney agenesis. The levels of Six2 are reduced in Six1 mutants, suggesting Six1 acts upstream of Six2 (Xu et al., 2003; Li et al., 2003). Clearly, although not essential for activation of Six2, Six1 might play a role in establishing normal Six2 levels prior to the termination of Six1 expression around E11.5 (Xu et al., 2003). By that time, Six2 is thought to regulate its own activity through auto-feedback loops mediated by proximal and distal enhancer elements (Brodbbeck et al., 2004; Gong et al., 2007; Park et al., 2012). Collectively, these studies demonstrate quite distinct temporal expression patterns and regulatory dynamics for Six1 and Six2 in mouse kidney development.

Many of the genes integral for mouse kidney development are associated with renal anomalies in the human population, suggesting close genetic parallels between the two species. Mutations have been identified in a number of genes encoding transcription factors, signaling proteins, and receptors that act within the nephron progenitor niche or the adjacent ureteric epithelium, including EYA1, PAX2, SALL1, RET, BMP4, FGF20, ITGA8, and SIX1 and SIX2 (Müller et al., 1997; Davidson, 2009; Cain et al., 2010; Barak et al., 2012; Humbert et al., 2014).

SIX1 mutations are associated with branchio-oto-renal (BOR) syndrome, whereas SIX2 mutations are linked to isolated cases of renal hypodysplasia (Ruf et al., 2004; Weber et al., 2008), highlighting their crucial roles in human kidney development. Furthermore, SIX1 and SIX2 mutations have also recently been associated with Wilms’ tumor, a pediatric kidney cancer (Wegert et al., 2015; Walz et al., 2015). The tumors are characterized by blastemal, epithelial and stromal elements much like the developing kidney. The blastema displays nephron progenitor-like characteristics, expressing factors such as CITED1, SIX1 and SIX2 (Li et al., 2002; Lovvorn et al., 2007; Murphy et al., 2012; Sehic et al., 2012, 2014). Mutations in the DNA binding homeodomain of SIX1 and SIX2 are associated with chemotherapy-resistant blastemas, suggesting that these mutations might contribute to an aggressive etiology of such tumors (Wegert et al., 2015).

Although genetic studies support a common set of regulatory factors underlying mouse and human kidney development, there is clearly a marked difference between their nephron progenitor programs. Whereas the mouse kidney generates around 13,000 nephrons over approximately 2 weeks of active nephrogenesis, the human kidney forms around a million nephrons over a 30 week period of nephrogenesis (Cullen-McEwen, et al., 2003; Bertram et al., 2011). These striking differences between the duration and output of the nephron progenitor pool between mouse and man are likely to reflect different regulatory properties intrinsic to the progenitor pool or within the niche where progenitors reside.

In this study, we explored the intrinsic regulatory programs at play within human nephron progenitors and provide evidence for distinct regulatory programs of Six/SIX between mouse and human kidneys. The data provide a potential mechanistic link to the lengthened period of progenitor self-renewal and nephrogenesis underlying human kidney development.

RESULTS

Given the crucial role for Six2 in mouse nephron progenitor self-renewal (Self et al., 2006), our previous analysis of Six2-directed regulatory circuitry in nephron progenitors (Park et al., 2012) and the contribution of SIX2 mutations to human renal anomalies (Weber et al., 2008), we examined SIX2 regulatory function in the human fetal kidney. Six2 is highly expressed from E10.5 within the mouse nephron progenitor population, and downregulated upon commitment of progenitors to nephron formation (Oliver et al., 1995; Self et al., 2006; Kobayashi et al., 2008; Mugford et al., 2009; Park et al., 2012) (Fig. 1A). Human kidney development initiates at 5 weeks with the invasion of the ureteric bud and terminates around 36 weeks. Consequently, the 16 week human kidney is approximately one-third of the way through the active period of nephrogenesis, analogous to the E15.5-E16.5 mouse kidney, a developmental stage extensively characterized for Six2 regulation in earlier studies (Park et al., 2012; Kanda et al., 2014). Additionally, the kidney appears to be undergoing active branching until at least 20 weeks of gestation (L.L.O. and A.P.M., unpublished observations).

As in the mouse kidney, human SIX2 displayed a nuclear localization within condensed mesenchyme cells surrounding the ureteric epithelial tips in the outer kidney cortex (Fig. 1A). CITED1, a definitive nephron progenitor marker in the mouse (Boyle et al., 2007; Park et al., 2012), colococalized with SIX2 in this group of cells (Fig. S1A). The overlap of SIX2 and CITED1 was observed in all nephron progenitors, but unlike the mouse, where only Six2 extends into early stages of nephrogenesis, we observed human CITED1 beneath the ureteric branch tips in what are likely to be early-forming nephron structures (Mugford et al., 2009; Park et al., 2012; Fig. S1A). SIX2 activity extended into nascent nephron precursors underneath the ureteric buds. SIX2 expression was downregulated in the differentiating structures and localized proximally in the renal vesicle, which were both similar attributes to mouse SIX2 expression (Fig. 1A, bottom panel). Thus, the overall distribution of mouse and human Six2/SIX2 are quite similar, consistent with SIX2 highlighting the human nephron progenitor compartment. Unlike mouse kidneys, human kidneys have an underlying lobular organization. Where the lobes ingress and meet, SIX2+ progenitor niches closely abut each other but appear to maintain their local tip niche integrity with SIX2+ cells closely opposed to tips of the underlying branching ureteric tree (Fig. 1A, zoomed inset).

Next, we performed ChIP-seq on mouse and human kidney tissues to compare regulatory patterns between Six2/SIX2 and identify common and unique transcriptional targets. Human SIX2 binding was analyzed from two independent replicates of 17 week fetal kidney tissues. The QuEST ChIP-seq peak caller (Valouev et al., 2008) identified 54,068 and 1916 peaks, with a highly significant overlap of 1592 shared peaks between the two SIX2 datasets (P-value=10^{-43}; Fig. S1B). The differing number of peaks was due to lower levels of SIX2 ChIP enrichment in the second replicate (Fig. S1C). Examination of SIX2 binding near MEOX1 and WTI highlighted the similar binding profiles for SIX2 replicates (Fig. 1B). Within the mouse embryonic kidney MEOX1 is localized and restricted to nephron progenitors (Mugford et al., 2009). WTI is expressed more broadly including the progenitors, (Mugford et al., 2009) and is essential for progenitor maintenance (Kreidberg et al., 2011). These striking differences between the duration and output of the nephron progenitor pool between mouse and man are likely to reflect different regulatory properties intrinsic to the progenitor pool or within the niche where progenitors reside.
1993). SIX2 ChIP-seq peaks tended to localize within conserved blocks of DNA consistent with SIX2 binding to the conserved cis-regulatory elements (Fig. 1B). Multiple sites of SIX2 binding are found within WT1 introns and 100’s of kilobases (kb) 5’ and 3’ of the WT1 transcription unit, suggesting that WT1 is a major target of SIX2 regulation (Fig. 1B, bottom panel). Given that the reads for the
two replicates were more strongly correlated when enrichment was compared within rep1 binding regions ($R^2=0.46$ versus 0.19, Fig. S1C), indicating that SIX2-rep1 is a considerably stronger dataset, we restricted further analyses to SIX2-rep1.

Approximately 60% of SIX2 peaks mapped within 50-500 kb of transcriptional start sites (TSSs) (47% randomly expected, $P$-value=$10^{-129}$); very few (<5%) were observed within 5 kb of the promoter (Fig. 1C). Additionally, SIX2 peaks predominantly occurred within intergenic (48%) and intronic (46%) regions, which is a typical pattern of bona fide enhancers (Fig. 1D). We performed a motif search within ±100 bps of the center of the top 1000 SIX2 peaks using the de novo motif finder MEME (Bailey et al., 2009). The top motif, TCANGTTTCA, closely matches a previously verified SIX2 binding motif from FACS sorted nephron progenitors (Park et al., 2012) that mapped to 60% of all SIX2 peaks (Fig. 1E). Motifs were enriched at the peak center as expected for a direct association of SIX2 with DNA (Fig. 1F). Furthermore, calculating the average conservation PhyloP scores (Siepel et al., 2006) across motif bases within SIX2 peaks demonstrated that the high-frequency motif bases tended to also have higher conservation (Fig. 1E). These data highlight the functional significance in the conservation of nucleotides that are likely to mediate DNA-protein contacts (Kumar, 2009).

To further interrogate the biological functions of human SIX2, we performed GREAT GO analysis (McLean et al., 2010) on the ChIP-seq peaks. SIX2 peaks were highly enriched near genes associated with metanephric kidney specific processes such as ‘nephron morphogenesis’ and ‘metanephric development’, and predicted a target cell type with an appropriate ‘metanephric mesenchyme’ and ‘urogenital system’ gene expression signature (Fig. 1G). In summary, analysis of human SIX2 ChIP-seq data uncovers a robust set of SIX2-bound enhancers within human nephron progenitors supporting a role for SIX2 regulation of nephron progenitor programs in the progenitor niche.

To assess the potential functional similarities and differences between human SIX2 and mouse Six2, we compared human ChIP-seq data with an E16.5 mouse whole kidney Six2 ChIP-seq dataset. The mouse data recovered an identical SIX2 binding motif to that of the human SIX2 ChIP-seq data. Similar to human SIX2, a large fraction (43%) of mouse Six2 peaks contained a Six2 motif enriched at peak centers (Fig. 2A,B). Because the mouse and human datasets were roughly comparable in their strength, we used a stricter peak-calling threshold to identify the strongest set of peaks: 12,145 for the mouse kidney and 6276 for the human kidney. In order to compare binding patterns of SIX2/Six2 between the two species, we ‘humanized’ mouse Six2 peaks by converting mouse peak coordinates to their human counterparts with the UCSC genome browser liftOver tool (Rhead et al., 2010).

Of the 9004 converted mouse Six2 peaks, only 727 sites (~8%) overlapped with human SIX2 peaks with a gap threshold of 100 bp (Fig. 2C). The small degree of peak overlap cannot be attributed to differences in the antibodies used for the human and mouse ChIP-seq comparison, because the two antibodies produce correlated binding data in mouse (Fig. S2A). Reproducible peaks were enriched for kidney target genes and the Six2 motif (Fig. S2B,D) and differential peaks tended to localize close to the TSS of highly active metabolic genes, without kidney specificity and were not enriched for the Six2 motif (Fig S2B,C,D). The finding of low binding site overlap between mouse and human is in line with previous reports comparing transcription factor binding in the same cell or tissue between species (Odom et al., 2007; Kunarso et al., 2010; Schmidt et al., 2010). For example, only 12-14% of the binding sites for CEBPa and HNF4A in the mouse and human liver are conserved; the differences have been attributed to the loss of motifs as a result of sequence changes (Schmidt et al., 2010). SIX2 human/mouse shared sites show the greatest enrichment for the SIX2 motif, 65% compared with 59% (human unique) and 21% (mouse unique), and the strongest conservation of the recovered binding motif (Fig. 2C). Furthermore, shared peaks had better enrichment of GO terms associated with kidney function such as ‘urogenital system development’ and ‘metanephros development’. These terms were absent from unique peak sets (Table S2). These observations argue that shared mouse-human sites have stronger functional roles compared with peaks observed in only one species. Interestingly, despite a relatively small overlap of Six2 binding sites between mouse and human, over 50% of putative Six2 target genes are shared between the two species (Fig. 2D). These results support the idea that Six2 binding is more conserved at the level of target genes, compared with conserved binding at individual enhancers. Therefore, new Six2 sites have probably evolved near the same target genes, contributing to regulatory and species diversity.

As suggested by the GO analysis, the overlap of mouse and human binding sites was enriched for potential target genes associated with kidney functions (Table S2). This includes genes that have integral roles in mouse kidney development and associate with human renal abnormalities, such as SALL1, EYA1 and SIX2 (Abdelhaky et al., 1997; Kohlhase et al., 1998; Xu et al., 1999; Nishinakamura et al., 2001; Self et al., 2006; Weber et al., 2008). Our previous study showed that Eya1 is a direct target of Six2 (Park et al., 2012) acting through an enhancer that is also conserved and bound by human SIX2 (asterisk in Fig. S3). Additionally, several other potential enhancer modules around Eya1/Eya1 are conserved between the two species (Fig. S3). These data highlight the conservation of cis-regulatory modules around genes with important roles in kidney development.

To discover potential novel Six2/SIX2 targets in mouse and human nephron progenitors, we utilized a combination of target regulatory potential and expression data. The regulatory potential measure is based on the number of peaks near each gene and the strength of the peaks (Tang et al., 2011). We first set out to identify genes with marked disparity in SIX2 regulatory potential between mouse and human nephron progenitors (Fig. 2E, left panel, Tables S2 and S3). As expected, SIX2 is a strong putative target of its own regulation in both mouse and human (Park et al., 2012; Fig. 2E, Tables S2 and S3). Surprisingly, one of the most highly regulated targets of human SIX2 was SIX1 (Fig. 2E, Tables S2 and S3). In the mouse, SIX1 expression is lost shortly after SIX2 is turned on (Xu et al., 2003) and therefore is an unlikely target. In agreement with this, SIX1 had the lowest possible regulatory potential in mouse, as expected from its temporally restricted expression profile (Xu et al., 2003; Fig. 2E, Tables S2 and S3). These data also indicated that Six2 is not likely to directly repress SIX1. Thus, SIX1 appears as a human-specific target by analysis of regulatory potential.

To further narrow down the list of genes identified as species-specific targets by regulatory potential, we examined their expression in human and mouse nephron progenitors to identify targets that also have species-specific expression. We performed RNA-seq on FACS isolated ITGA8<sup>+</sup> cells from 17 week human fetal kidney cortex and E15.5 Cited1<sup>+</sup> nephron progenitors. ITGA8 is expressed in the nephron progenitors and induced structures of the kidney (Müller et al., 1997; Fig. S4A). We utilized a limited enzymatic digestion of the human fetal kidney to isolate cells from the outer cortical layers in a procedure that recovers ITGA8<sup>+</sup>...
nephron progenitors but excludes the majority of differentiating structures (Fig. S4A). Using RNA-seq data from nephron progenitors, we compared expression of genes between human and mouse and identified genes that were >5-fold enriched in either species and were also a species-specific target (Fig. 2E, right panel, highlighted genes).

*SIX1* is expressed in the human ITGA8+ progenitors, but not in mouse nephron progenitors, identifying *SIX1* as a human specific target on the basis of both cis-interactions around the *SIX1* gene and active *SIX1* expression (right panel Fig. 2E; Table S3).

We examined epigenetic chromatin signatures around *SIX2/Six1* and *SIX2/Six2* genomic regions in both species to identify regulatory differences that might contribute to species differences in *SIX1/Six1* expression. ChIP-seq was performed on 17 week fetal kidneys and E16.5 mouse kidneys to assess chromatin marks associated with active genes and enhancers (H3K27ac) and transcriptionally silenced chromatin (H3K27me3). In the human fetal kidney, the *SIX2* locus displayed a similar profile to that of the mouse: bound by SIX2 at conserved elements and marked by H3K27ac in both the gene body and at SIX2-bound regions (Fig. 3B). Similarly, the human *SIX1* locus was bound by SIX2 at multiple conserved elements and displayed prominent H3K27ac throughout the gene body and the SIX2-bound regions (Fig. 3A). By contrast, the mouse *Six1* locus did not show significant binding by Six2 or H3K27ac enrichment but was marked by a strong H3K27me3 signal (Fig. 3A), which is consistent with epigenetic silencing of the region. Together, these findings indicate a transcriptionally active human *SIX1* state and suggest that *SIX1* expression might be regulated, at least in part, through direct SIX2-mediated transcriptional activation.

Previous immunodetection studies reported SIX1 localization in the condensed mesenchyme of the 17-20 week human fetal kidney (Li et al., 2002; Sehic et al., 2012). However, SIX1 and SIX2 are highly conserved in their DNA-binding and SIX domains (Fig. S4C); consequently, the potential for crossreactivity of antibodies between SIX proteins clouds this interpretation. To definitively examine SIX1 localization in the developing human kidney, we utilized a C-terminal-specific antibody that uniquely recognizes SIX1 (Fig. S4B). At 16 weeks of fetal development, nuclear SIX1 was readily identified within nephron progenitors throughout the many nephron progenitor niches established following the onset of ureteric branching 11 weeks earlier (Fig. 4A). Furthermore, SIX1 and SIX2 proteins showed a highly similar distribution in human nephron progenitors (Fig. 4A). By contrast, mouse *Six1* and *Six2* overlapped in the metanephric mesenchyme at E10.5, but *Six1* was absent from nephron progenitors by E11.5 (Fig. 4B). However, Xu et al. (2003) observed *Six1* activity at E11.5 through a lacZ knock-in allele. Because *Six1* activity was measured indirectly, this finding is
likely to reflect perdurance of β-galactosidase activity following silencing of Six1.

Because Six1 and Six2 are transiently co-expressed in the E10.5 metanephric mesenchyme, we asked whether Six1 expression was dependent on Six2 at this stage. We examined Six1 expression in Six2GCE/+ and Six2GCE/GCE mouse kidneys that harbor a mutant allele generated by knock-in of a GFP cassette into the Six2 locus (Kobayashi et al., 2008). In both heterozygous and Six2-null mutants, we observed co-labeling of GFP and Six1 in the metanephric mesenchyme, with similar levels of Six1 staining between the two genotypes (Fig. 4C). Therefore, Six1 activity is not dependent on Six2 in the metanephric anlagen. Furthermore,
Although all GFP+ cells were SIX1+, there were many more cells that had Six1 expression but lacked GFP signal. This suggests that Six1 activation precedes and is independent of Six2, consistent with a requirement for Six1 in the E10.5 kidney and the more severe Six1 mutant phenotype (Xu et al., 2003; Li et al., 2003; Self et al., 2006; Xu and Xu, 2015).

The SIX2-bound regions near the human SIX1 locus might serve as enhancers maintaining SIX1 expression in the human fetal kidney. To examine the regulatory activity of these regions, we selected the two strongest SIX2-bound conserved modules within the SIX1 locus (Fig. 3A, asterisks) and tested a single copy of each for enhancer activity in a G0 mouse transgenic assay scoring for activation of a lacZ::nGFP fusion cassette. The strongest enhancer (Enh1) lies in an intergenic region ∼11.5 kb downstream of the SIX1 promoter, and displays high conservation across vertebrates (Fig. 3A). The second strongest enhancer (Enh2) lies ∼4 kb upstream of the SIX1 promoter within another highly conserved block (Fig. 3A). These two enhancers were previously confirmed to be regulatory elements controlling Six1 expression in the developing mouse embryo (Sato et al., 2012). Enh1 and Enh2 both showed activity in the otic vesicle and cranial ganglia, reported sites of Six1 expression (Sato et al., 2012) (Fig. 5A).Enh1 showed additional activity in the olfactory placode, eye and apical ectodermal ridge of the developing limb bud. However, only Enh2 showed highly reproducible metanephric mesenchyme-specific expression at E10.5 (0/3 for Enh1, 6/9 for Enh2; Fig. 5A), similar to their mouse equivalents (Sato et al., 2012). When Enh1 was analyzed at E15.5, 1/17 transgenic positive kidney pairs displayed a nephron progenitor-specific expression pattern, whereas 3/17 displayed additional distinct β-gal+ patterns (Fig. 5B,C; Fig. S5). For Enh2,
1/5 transgenic kidney pairs showed a mosaic expression that mapped specifically to nephron progenitors (Fig. 5B,C). In summary, only Enh2 showed robust activity in the mouse metanephric mesenchyme at E10.5, whereas both enhancers showed sporadic nephron progenitor activity at E15.5 when mouse Six1 was inactive. Collectively, these data highlight early active enhancer elements switched on prior to active nephrogenesis that are mostly, but not always, shut down in later nephron progenitors (see Discussion).

To directly address the functional role of SIX1 in the human fetal kidney, we performed SIX1 ChIP-seq on 16 and 17 week kidney replicates. The two datasets showed moderate overlap and were correlated (Fig. S6A,B). To remove potential false-positive SIX1 peaks, we focused on the overlapping set of 1610 sites. De novo motif recovery identified a peak-centered motif matching the SIX2 motif, consistent with the highly conserved DNA-binding homeodomain of SIX1 and SIX2 and previous SIX1 ChIP data from C2C12 myoblast cells (Liu et al., 2012; Fig. 6A,B; Figs S4C and S6C). The 1610 overlapping SIX1 peaks had a SIX motif recovery rate of 60%, similar to the SIX2 peaks (Fig. 6A) and higher than each individual SIX1 ChIP-seq replicate (43% and 38%, Fig. S6A), supporting the specificity of the shared SIX1 peaks and indicating that the strongest peaks within each dataset lie within the overlap. Thus, SIX1 and SIX2 recognize the same DNA binding motif and consequently, each factor is likely to target a common set of enhancers and regulate a common set of genes in the nephron progenitor pool.

Consistent with this prediction, an overwhelming majority of SIX1 peaks (~81%) were shared with SIX2 peaks, and their binding strengths were significantly correlated (R²=0.4; Fig. 6C,E). Additionally, nearly all predicted SIX1 target genes (~90%) were shared with SIX2 (Fig. 6D). SIX1-only peaks had lower signals compared with shared peaks (Fig. S6E), indicating that they represent peaks where SIX2 signals fall below the detection threshold rather than being truly unique sites. The C-terminal regions of SIX1/2 protein sequences are divergent and could lead to...
differing protein-protein interactions (Fig. S4C), which might influence levels of SIX1 and SIX2 recruitment to their target sites through interactions with differing co-factors. This idea is supported by a relatively low correlation (0.18) between SIX1 and SIX2 signals across the shared peaks (Fig. S6D). Motif recovery on the shared and non-overlapping peaks identified a WT1-like motif and E-box motif in both the overlapping and SIX2 only sites (Fig. S6F), suggesting that WT1 and a bHLH factor are potential binding partners of SIX1 and SIX2. The 303 SIX1-only sites yielded a SIX motif, but no WT1 or E-box signals. The lack of co-factor motifs amongst Six1-only sites is most likely due to the low number of peaks and low enrichment of these peaks (Fig. S6E,F). Whether SIX1 and SIX2 interact with different co-factors at independent target sites remains an open question.

Fig. 6. SIX1 and SIX2 share common targets and show evidence of auto- and cross-regulatory activity. (A) Comparison of the most enriched motif for SIX1 and SIX2 peaks. (B) Distribution of SIX1 motif-peak distances. (C) Overlap of SIX1 and SIX2 binding sites. (D) Overlap of SIX1 and SIX2 target genes. (E) Comparison of raw signals from SIX2 and SIX1 ChIP-Seq data sets. Each point represents a single binding peak. (F) Gene ontology analysis of shared SIX1/SIX2 peaks. Obs., observed; Exp., expected. (G) Genomic view of the human SIX1 (left) and SIX2 (right) gene loci. (H) Western blot of SIX2-3×FLAG co-immunoprecipitations from HEK293 cells. (I) Western blot of SIX1 and SIX2 co-immunoprecipitations from human fetal kidneys.

RESEARCH ARTICLE
GREAT GO analysis of the overlapping SIX1-SIX2 peaks revealed an association with kidney processes such as ‘metanephros development’ and expression of the targets in kidney associated structures such as ‘metanephric mesenchyme’ (Fig. 6F). Predicted target genes include SIX1, SIX2, SALL1, WT1 and OSR1 (Table S4). Taken together, these data indicate that SIX1 and SIX2 recognize a very similar set of enhancers for the same targets in human nephron progenitors mediated through interactions with a common SIX-type motif. Importantly, these interactions include co-regulatory inputs at their own and each other’s enhancers (Fig. 6G).

These findings leave open the possibility that both factors are simultaneously engaged within a common regulatory complex. To address this, we performed co-immunoprecipitations from HEK293 cells transfected with tagged proteins. First, as a positive control, we confirmed that SIX1 and SIX2 complex with EYA1 (Fig. 6H, data not shown), in agreement with previous studies using fly and mouse homologs (Pignoni et al., 1997; Buller et al., 2001). Next, we analyzed whether SIX1 and SIX2 interact with each other. Whereas, SIX1 and SIX2 were co-immunoprecipitated using specific antibodies for distinct epitope tags following overexpression in HEK293 cells (Fig. 6H, SIX1 data not shown), SIX1 and SIX2 were not co-immunoprecipitated by SIX1- and SIX2-specific antibodies in extracts of 17 week human kidney (Fig. 6I). The results indicate that either (1) the antibodies used for immunoprecipitation in vivo disrupt the heterodimeric SIX1-SIX2 complex or (2) SIX1 and SIX2 form independent transcriptional complexes in vivo that are capable of associating with the same regulatory elements in human nephron progenitors, but ectopically, SIX1 and SIX2 can form complexes when present at high levels in a heterologous cell type. To distinguish between these possibilities, we repeated the co-immunoprecipitation analysis in HEK293 cells using the SIX2-specific antibody from the in vivo studies. In HEK293 cells, SIX2 and SIX1 co-immunoprecipitated indicating the SIX2-specific antibody does not disrupt the in vitro generated SIX1-SIX2 complex (Fig. 6H). Taken together, these results support the presence of independent SIX1 and SIX2 regulatory complexes in vivo, although we cannot rule out the possibility of some minor role for less-stable SIX1-SIX2 complexes that might be highlighted by in vitro overexpression conditions.

DISCUSSION

Comparison of human and mouse SIX2/Six2 functions

In this study we examined the conservation of human and mouse SIX2/Six2 expression and function. We observed that localization of SIX2 within nephron progenitors is similar in the developing mouse and human fetal kidney. A majority of the SIX2/Six2 transcriptional targets are shared between the two species, demonstrating a common set of SIX2/Six2 target genes despite a relatively low overlap of binding peaks at the homologous enhancers. Additionally, the in vivo recovered motif bound by Six2/SIX2 is identical in mouse and human progenitors, in agreement with conserved DNA-binding domains. Given that their Six/SIX domains and C-terminal domains are also highly similar, protein-protein interactions mediated through these regions are also likely to be conserved between the two species.

A recent study has shown that conserved sites bound by transcription factors in mouse and human are correlated with pleiotropic functions (Cheng et al., 2014). These enhancers are active across multiple tissues, subjecting them to strong evolutionary constraints that preserve motifs within enhancer modules. The authors suggest that the conserved, pleiotropic enhancers might be bound by transcription factors within the same family that recognize the same motif (Cheng et al., 2014). In our data, the conservation of the SIX2 motif is highest amongst shared sites of human and mouse binding and enhancers are conserved around target genes such as EYA1/Eya1. Eya1 is integral for the proper development of several tissues, including the kidney, inner ear, cranial ganglia and branchial arch derivatives (Xu et al., 1999, 2002; Zou et al., 2004). It would be interesting to determine whether our prospective enhancers are also active in these additional tissues. SIX1 and SIX2 are expressed in many of these same tissues, consistent with a multi-tissue regulatory link (Oliver et al., 1995; Sato et al., 2012).

Whereas the target genes and function of SIX2 appear to be highly conserved, we identified SIX1 as a novel and unique target of SIX2 in the human fetal kidney. Our analyses uncovered other gene targets predicted through regulatory potential and expression analyses to show species-specific patterns of progenitor activity. Other than SIX1, our data identifies several additional genes that have high expression in the human kidney ITGA8+ cells versus mouse nephron progenitors, and have higher regulatory potential in the human versus mouse (Table S3). Similar to SIX1, such genes represent unique regulatory targets of SIX2 in human. Examples of such genes include COL6A2 and CDH7, suggesting potential differences in matrix and cell-cell adhesions between human and mouse nephron progenitors. Conversely, Hs3st6, a heparin sulfate sulfotransferase and Wt1 target (Motamedi et al., 2014), represents a mouse-specific SIX2 target gene with higher expression in mouse Cited1+ nephron progenitors (RPKM=35.640) but low expression (RPKM=0.225) in the ITGA8+ human nephron progenitor-enriched population. Confirmation of species-specific expression of these genes and their potential differential impact on mouse and human nephron progenitor functions will be a focus for future studies.

SIX1 function in mouse versus human

SIX1 is required for maintenance of the early metanephric mesenchyme (Xu et al., 2003; Li et al., 2003; Xu and Xu, 2015), but by the time the first round of branching has occurred in the mouse, Six1 is no longer detectable. However, SIX1 activity extends far beyond the initial round of branching, and overlaps with SIX2 in human nephron progenitors. These findings raise the questions of (1) how mouse and human differentially regulate their Six-genes during kidney development, and (2) what is the functional significance of their divergent regulatory programs?

Clearly, a common regulatory theme for mouse Six2 and human SIX1/2 are their auto-regulatory activities. Each factor binds its own gene’s progenitor-specific enhancer; in addition, human SIX1 and SIX2 cross-regulate SIX2 and SIX1 genes, respectively. However, their initial activation in the early-specified metanephric anlagen is likely to be dependent on other factors. In the mouse, our data demonstrate that Six1 activation is independent of Six2 and that Six1 acts upstream of Six2, in line with previous reports showing that Six1 is required for normal Six2 expression (Xu et al., 2003; Li et al., 2003). The situation in the equivalent stage of human kidney development (4.5-5 weeks) is presently unknown (Fig. 7).

Examination of transgenic activity of human SIX1 (this paper) and mouse Six1 (Sato et al., 2012) enhancer modules suggests that initial activating mechanisms might be regulated through a common enhancer, and this module and potentially others, promotes persistent SIX factor-mediated nephron progenitor expression of human SIX1. In this scenario, enhancer silencing within the mouse Six1 locus through activities of additional
regulators would presumably block engagement of Six1 and Six2, resulting in early down-regulation of Six1 in the mouse. In the human kidney, such enhancer silencing activities are absent and both SIX1 and SIX2 expression persists in nephron progenitors through weeks of highly active nephrogenesis (Fig. 7). Alternatively, human SIX1 might utilize distinct regulatory elements not shared with the mouse and excluded from Enh2 that maintain SIX1 expression in human nephron progenitors after the initial activating trigger is lost.

The human SIX1 enhancers show the most robust and consistent activity at E10.5, with Enh2 displaying metanephric mesenchyme activity at E10.5, overlapping endogenous Six1 expression (Xu et al., 2003). However, their reporter expression patterns become variable with rare activity by E15.5. Whereas over 50% of transgenic mice show activity from a Six2 distal enhancer (Park et al., 2012; L.L.O. and A.P.M., unpublished data), only 6-20% show activity from the human enhancers at this time. This suggests that human enhancers are subject to similar regulation to their mouse Six1 regulatory counterparts but might escape that regulation in some transgenic lines where the transgene integration site could influence the expression outcome. Importantly, as SIX1/2 binding motifs are conserved between mouse and human in both enhancers (Fig. 3A), the differing regulatory outcomes for mouse Six1 and human SIX1 do not appear to result from the loss of Six-specific binding elements.

In mouse nephron progenitors at E16.5, the Six1 locus is marked by an H3K27me3 signature indicative of PRC2-mediated transcriptional silencing. When and how this silencing occurs remains to be determined. In the mouse, Six2 progenitor expression extends until depletion of the nephron progenitors at the end of the nephrogenic period (Hartman et al., 2007; Rumballe et al., 2011). The temporal expression patterns of SIX1 and SIX2 throughout human nephrogenesis are currently unclear. SIX2 expression in human fetal kidney progenitors has been reported at 24 weeks of development but nephrogenesis continues until 36 weeks (Murphy et al., 2012).

The functional significance of distinct Six1/SIX1 regulation between mouse and human is a matter for speculation at this time. Clearly, Six1 and Six2 are key regulators of nephron progenitors (Xu et al., 2003; Li et al., 2003; Self et al., 2006; Xu and Xu, 2015), and Six2 maintains and expands progenitors by countering progenitor commitment to nephrogenesis (Self et al., 2006; Kobayashi et al., 2008; Park et al., 2012). One attractive model posits a dual action for SIX1 and SIX2 in modifying progenitor programs to extend the period of progenitor expansion. Overexpression of SIX2 in a nephroblastoma cell line increases the number of cells in S-phase (Senanayake et al., 2013), supporting the idea that elevated levels of SIX proteins enhance cellular proliferation and progenitor expansion. SIX2 is expressed at higher levels than SIX1 in the ITGA8+ progenitor-enriched population (157.12 RPKM versus 21.57 RPKM, respectively; Table S3) suggesting that SIX2 remains the predominant SIX factor in nephron progenitors. A relatively small change in SIX levels could have significant ramifications in the balance of progenitor numbers. Furthermore, recent evidence suggests that OSR1/Osr1, which shows comparable expression levels to SIX1 in both human and mouse (RPKMs of 16.21 and 32.01, respectively; Table S3), acts synergistically with Six2 to maintain the nephron progenitors (Xu et al., 2014). Additional experimental studies will be required to explore the significance of human SIX1 in expanding nephron progenitors.

Six2 and its human SIX counterparts also bind enhancers that activate expression of genes encoding key signals promoting progenitor differentiation such as Fgf8 and Wnt4, suggesting a role for Six/SIX factors in the control of progenitor commitment (Park et al., 2012 and data therein). Ultimately, the period of progenitor activity depends on a balance of progenitor renewal and commitment, altering the dynamics of either process will influence the size of the nephron progenitor pool and the duration of nephrogenesis. The extended lifetime of the human nephron progenitor pool is likely to be a key factor in the 100-fold greater number of nephrons formed in the human versus the mouse kidney. Further mechanistic insights might be gained from examining regulation of Six1/2 and regulatory activity in experimental mammalian models with a nephrogenic period and nephron count closer to the human kidney.

Fig. 7. Model of differential regulation of Six1/2 and SIX1/2 in the developing mouse and human kidney. In the metanephric mesenchyme of the mouse (E10.5), Six1 expression is driven by factor (s) Xc and is actively transcribed (Pol II). Six1 can then activate Six2 expression (1), and subsequently Six2 can drive its own expression (2) via an autoregulatory loop. Because both loci are active, they are marked by H3K27ac (ac). However, in the mature nephron progenitors, Six1 is no longer expressed and displays a repressive histone signature of H3K27me3 (me3). Six2 cannot access the Six1 enhancers and continues to drive its own expression. In the human metanephric mesenchyme (~5 weeks of gestation), the expression and regulation of SIX1 and SIX2 are unknown. In mature nephron progenitors and in contrast to the mouse, SIX1 is active and expression is driven by SIX2 and itself. Similarly, SIX2 expression is driven by SIX1 and itself. SIX1 and SIX2 are likely to regulate expression through discrete complexes.
SIX1/2 in cell programming and Wilms’s tumor

Recent reports have shown that SIX1 and SIX2 are both required to reprogram human proximal tubule cells to nephron progenitors (Hendry et al., 2013), suggesting that SIX1 might have additional non-overlapping functions with SIX2. Alternatively, absolute levels of SIX proteins might be important, and high levels are required for reprogramming and progenitor maintenance. SIX1/2 mutations have recently been associated with Wilms’ tumors (Wegert et al., 2015; Walz et al., 2015). These data provide further evidence that the blastemal elements of the tumor reflect the characteristics of the nephron progenitor niche. Interestingly, Wegert et al. (2015) performed SIX1 ChIP-seq on tumor samples with and without the SIX1 mutation. The motif recovered from the wild-type tumor type, GAAACCTGATCC, closely matches the TGAAACCTGA recovered from the SIX1/2 motif. A comparative analysis of tumor and developmental programs might identify regulatory networks and gene targets responsible for the persistence of these tumor cells through chemotherapy treatment.

MATERIALS AND METHODS

Mouse and human kidney samples

All surgical procedures, mouse handling and husbandry were performed according to guidelines issued by the Institutional Animal Care and Use Committees (IACUC) at the University of Southern California and after approval from the institutional IACUC committee. Mouse strains utilized are described in supplementary Materials and Methods. De-identified human fetal kidney tissues ranging from 16-17 weeks gestation were obtained from Novogenix Laboratories following informed consent and elective termination. Developmental age was determined by ultrasound.

ChIP-seq

ChIP-seq from E16.5 mouse kidney tissue was performed essentially as described (Park et al., 2012). For ChIP-seq from human fetal kidneys, samples were microdissected to remove the cortex and incubated for 30 min at room temperature in crosslink buffer (Park et al., 2012). Crosslinking was stopped by the addition of glycine. Tissue was washed with PBS containing protease inhibitors (PI), homogenized, pelleted, and lysed in mouse ChIP lysis buffer with the aid of a B Dounce homogenizer. Processing of the samples from this point was carried out using the mouse ChIP protocol. ChIP was performed with antibodies listed in Table S1. Sequencing libraries for both mouse and human ChIP DNA were made using the ThruPLEX-FD Prep Kit (Rubicon Genomics). Libraries were sequenced at the USC Epigenome Center on the Illumina HiSeq 2000.

Fluorescence activated cell sorting (FACS)

E15.5 kidneys from Cited1-nuc-TagRFP-T+embryos were isolated and processed for FACS as described (Park et al., 2012). To isolate ITGA8+ cells from human fetal kidneys, the outer capsule was removed and kidneys incubated with Liberase (Roche) to remove the outer cortical cell layers. Cell suspensions were incubated with anti-ITGA8 (R&D, AF4076) and appropriate Alexa Fluor-labeled secondary antibody. Sorting was performed on a BD FACSARia II Flow Cytometer.

RNA-seq

RNA was isolated from Cited1-/- and ITGA8+ cells using the Qiagen RNeasy Micro Kit. RNA was submitted to the USC Epigenome Center for library preparation and sequencing on the Illumina HiSeq 2000. All RNA-Seq reads were aligned to hg19 or mm9 using the DNA Nexus and quantified to generate RPKM. RNA-seq data are available on the Gene Expression Omnibus (accession number GSE73867) and summary data are listed in Table S1.

Transgenic analysis of enhancer regions

G0 transgenic analysis was performed as previously described (Park et al., 2012). Details of enhancer construction and coordinates are described in the supplementary Materials and Methods. Samples were stained with X-gal, fixed and photographed using a Nikon SMZ 1500 fluorescent microscope.

For sections, stained kidneys were cryosectioned and immunostained as described below.

Immunofluorescence

16 week human fetal kidneys were fixed overnight. Mouse whole embryos or urogenital systems were fixed for 1 h. Human or mouse cryosections were immunostained as previously described (Mugford et al., 2008). Antibodies and dilutions are detailed in the supplementary Materials and Methods. Whole kidney images were captured on the Zeiss Axio Scan.Z1 Slide Scanner. All other images were acquired on a Nikon Eclipse 90i epi-fluorescent microscope, Zeiss LSM 780 inverted confocal microscope, or Leica TCS SP8 confocal. Slides from transfections were fixed and stained similarly using SIX1-specific, SIX2-specific or SIX2 crossreactive antibodies.

Immunoprecipitation analysis

HEK293 cells were transfected with pTARGET-SIX2-3*FLAG-P2A-mCherry, pCIG-SIX1-Mye, pCIG-EYAI-Mye, or the appropriate control empty vector. Details of construct generation and transfection can be found in the supplementary Materials and Methods. Nuclear lysates were prepared using the Active Motif Nuclear Complex Co-IP Kit. Extracts were incubated overnight at 4°C with anti-FLAG (F3165, SIGMA) antibody bound to Dynabeads Protein G (Life Technologies). Beads were washed six times following the Co-IP kit protocol with recommended high-stringency conditions. Samples were resolved on a 10% SDS-PAGE gel, transferred to nitrocellulose and subjected to standard western blotting protocols using anti-FLAG or anti-Myc antibodies. For tissue co-immunoprecipitations, the outer cortex of 17 week human fetal kidneys was microdissected and a nuclear lysate prepared. Normal rabbit IgG, SIX1 (Cell Signaling) or SIX2 (MyBioSource) antibodies were crosslinked with dimethyl pimelimidate to Dynabeads Protein G using the Protein A/G SpinTrap Buffer Kit (GE Healthcare). Nuclear extracts were incubated overnight with antibody cross-linked beads at 4°C. Samples were washed five times with TBS+0.1% Triton X-100 and proteins eluted with 0.1 M Glycine-HCl, pH 2.9. Samples were treated as above for western blots. Antibodies against SIX1 (1:1000) and SIX2 (1:1000) were used for detection.

ChIP-seq data analysis

All ChIP-seq sequences were mapped to hg19 or mm9 using Novoalign software (Novocraft). Mapped ChIP-seq and input data were analyzed using QuEST 2.4 software (Valouev et al., 2008). Mouse binding sites were converted to human sites using the liftOver utility (Rhead et al., 2010) available at the UCSC genome browser website. ChIP summary data is listed in Table S1. Human SIX1/SIX2 and mouse Six2 motifs were calculated using MEME de novo motif finder (Bailey et al., 2009). To assess the evolutionary conservation of the motif sites, we retrieved the cross-species “PhyloP” conservation scores from the UCSC genome browser (Siepel et al., 2006). GREAT GO analysis was performed using the online GREAT program v2.0 (McLean et al., 2010). DAVID (Huang et al., 2007) was used on target genes from the analysis human and mouse binding site overlap. Assignment of target genes was performed by associating peaks with genes using GREAT (McLean et al., 2010). More specific details of parameters used are described in the supplementary Materials and Methods. ChIP-seq data are available on the Gene Expression Omnibus under GSE73867.

Acknowledgements

The authors thank Dr Joo-Seop Park for helping develop the mouse whole kidney ChIP protocol and members of the McMahon Lab for critical discussion of the data. We thank Zayed Albertyn and Colin Hercus for their help with Novoalign.

Competing interests

The authors declare no competing or financial interests.

Author contributions

L.L.O. and A.P.M. designed the experiments. L.L.O. performed all experiments except the nephron progenitor RNA-seq (J.-D.B.), pronuclear injections (Y.L.), and...
HEK293 construct generation and transfection (Q.G.). T.T. helped with the human immunostaining and P.H.W. helped with the ITGA8 FACS. Q.G. and A.V. designed and performed bioinformatics analyses, L.L.O., Q.G., A.V. and A.P.M. analyzed the data. L.L.O., Q.G., A.V. and A.P.M. prepared the manuscript.

Funding
Work in A.P.M.’s laboratory was supported by grants from the National Institutes of Health [DK054364 and DK094526]. Q.G. was supported by a graduate student fellowship from the California Institute for Regenerative Medicine. Deposited in PMC for release after 12 months.

Supplementary information
Supplementary information available online at http://dev.biologists.orglookup/suppl?doi=10.1242/dev.127175/-/DC1

References


Supplemental Materials and Methods

Mouse strains

Six2<sup>GCE/+</sup> mouse lines were generated as previously described (Kobayashi et al., 2008) and maintained on a C57BL/6J (Jackson Laboratory) background. Wild type mice utilized for ChIP-seq and immunostaining were from Swiss Webster (Taconic) timed matings where noon on the day of vaginal plug detection is considered 0.5 days post coitum, or E0.5. Cited1-nuc-TagRFP-T<sup>Tg</sup> animals (http://gudmap.org) were maintained as a homozygous breeding stock and mated to Swiss Webster females for timed collections. Transgenic analyses were performed using the hybrid B6SJLF1/J (Jackson Laboratory) strain.

Transgenic analysis of enhancer regions

A single copy of each enhancer was inserted via Gateway cloning methods into a modified Hsp68-lacZ:<nGFP reporter construct (Tsanov et al. 2012). Enhancers were PCR amplified from genomic DNA or synthesized (Genewiz). Enh1 hg19 genomic coordinates: chr14:61,104,600-61,105,260. Enh2 hg19 genomic coordinates: chr14:61,119,745-61,119,994. Pronuclear injections were performed in house using standard protocols and injection of 2 ng/μl DNA into B6SJLF2 fertilized eggs.

Antibodies and dilutions used for immunostaining

SIX1 (1:500, Cell Signaling, 12891), SIX2 (1:1000, MyBioSource, MBS610128), CITED1 (1:250, Abnova, H00004435-M03 (Fig. S1A, Bottom); 1:500, Lab Vision, RB-9129 (Fig. S1A, Top)), JAG1 (1:250, R&D Systems, AF599), ECAD (1:250, BD Biosciences, 610181; 1:1000, Sigma, U3254), ITGA8 (1:500, R&D Systems, AF4076), GFP (1:500, AvesLabs, GFP-1020), cytokeratin (1:250, Sigma, C2931).

Cell culture and transfections

NIH3T3 or HEK293 cells (ATCC) were cultured in slide chambers or dishes, respectively, at 37°C, 5% CO₂ in DMEM containing 10% FBS, 100 units/ml penicillin and 100 mg/ml streptomycin. SIX2 was amplified from 16 week fetal kidney cDNA and cloned into the pCS2+ plasmid using BamHI and EcoRI restriction sites (SIX2-pCS2). SIX2 was amplified with a 3XFLAG tag from this construct and inserted into pTARGET carrying a -P2A-mCherry cassette using BamHI and XhoI restriction sites. SIX1-pCMV6-XL5 was purchased from OriGene. SIX1 was amplified with a Myc tag and inserted into the pCIG vector using XhoI and Clal sites. EYA1 was amplified from 16 week fetal kidney cDNA with a Myc tag and inserted into pCIG using EcoRI and Xmal sites. Empty vectors were used as controls. Transfections of all constructs were performed using Promega FuGENE HD transfection reagents following manufacturer protocols.
**ChIP-seq alignment and peak calling**

Novoalign (Novocraft) alignment parameters: Single-end reads trimming of 10 bp, polyclonal read filter: 7,10 0.4,2, maximum alignment score acceptable: 120. QuEST (Valouev et al., 2008) peak calling parameters: We used a “transcription factor” setting (for human SIX2, SIX1 and mouse Six2; bandwidth of 30 bp, regions size of 300 bp) or a ‘histone’ setting (for human H3K27ac, mouse H3K27me3 and H3K27ac; bandwidth of 100 bp, regions size of 1000 bp). Peak calling stringency was specified as following: a sample-dependent ChIP enrichment fold, 3-fold ChIP over input enrichment were used to seed the regions, and 3-fold ChIP enrichment was assigned for extending the regions. FDR for detecting the bound regions was evaluated by allocating the same number of mapped reads from a separate mouse input library and performing QuEST analysis using the same parameters.

**DNA sequence motif analysis**

MEME (Bailey et al., 2006) was run on a pool of 100 bp sequences obtained around the top 1000 most-enriched human SIX2, mouse Six2 or human SIX1 ChIP-seq peaks. Percentages of peaks with motif were calculated using FIMO tool (Grant et al., 2011) with p-value cutoff of 0.0002. We used binomial test to evaluate the statistical significance of motif enrichment within binding peaks.

**Calculation of regulatory scores**

In order to measure the potential of a gene being regulated by SIX2, we calculated regulatory score \( RS_g \) for a given gene \( g \) using the following formula:

\[
RS_g = \frac{\sum p \frac{I_p}{\max(D_p,20000)}}{\sqrt{p}},
\]

where \( p \) is the ID of the peaks associated with gene \( g \), and \( D_p \) is the distance from peak \( p \) to the TSS of gene \( g \). To select human/mouse-specific SIX2 target genes, regulatory scores were normalized to peak numbers and 75th quantile of peak intensity, then a z score was calculated based on Poisson distribution of regulatory score. Genes with \( z > 1 \) were selected as ‘human/mouse-specific genes by SIX2 regulation’. These genes were further filtered for those with \( z > 5 \) fold expression in human vs. mouse (or vice versa) to generate the ‘human/mouse-specific genes by both regulation and expression’.

**ChIP-Seq data comparison for mouse and human antibodies**

To compare the Six2 ChIP-Seq data generated using the Ab used in mouse Six2 ChIP-Seq (Six2_ms) and the in human SIX2 ChIP-Seq (Six2_hu), we generated 3 replicate ChIP-Seq data sets using each of the two Abs, respectively, in E16.5 mouse embryonic kidneys. After peak calling using QuEST, the peaks from all 3 replicates of each group were merged so that any peaks within 200 bp from each other are unioned, with the new coordinate being the midpoint of the original peak coordinates. Then ChIP-Seq
reads from each replicate were counted within +/-500 bp window around each peak. The read counts were transformed into enrichment fold (ChIP-Seq score, i.e. \( \frac{\text{reads in 1 kb window}}{1000} \div \frac{\text{total reads count}}{\text{genome size}} \)). We normalized the ChIP-Seq score for each replicate by bringing the median scores to the same number (practically, the maximum median). We performed pairwise comparison of the normalized data across replicates for Six2_ms and Six2_hu, respectively. To compare Six2_ms and Six2_hu data, we first merged the peaks from all six replicate data sets, and then calculated the ChIP-Seq score, which was subsequently normalized by bringing the medians of scores to the same number. Then we plotted the average ChIP-Seq score of Six2_ms vs. Six2_hu. Observing a subpopulation of peaks are skewed toward Six2_ms, we identified Six2_ms-specific sites by applying the threshold of Six2_ms score > 10 and Six2_cross score/Six2_spec score > 5. The reproducible peaks were defined as those with Six2_ms score > 10 and Six2_hu score > 10.

Protein alignment

Human and mouse SIX1/Six1 and SIX2/Six2 proteins were aligned using the ClustalW2 tool from the EMBL-EBI (Larkin et al., 2007). The SIX and Homeodomains were highlighted, and the amino acids that make DNA contacts marked (Kumar, 2009).
Figure S1: CITED1 expression is similar to SIX2 and comparison of human SIX2 ChIP-seq replicates indicate rep1 is a stronger dataset. (A) CITED1 and cytokeratin immunostaining of a 16 week human fetal kidney. Zoomed view of SIX2, CITED1, and ECAD co-staining shows overlap of SIX2 and CITED1 in nephron progenitors. (B) Venn diagram shows overlap between the two replicates of SIX2 ChIP-seq (SIX2-rep1, SIX2-rep2) peaks from independent 17 week human fetal kidneys. (C) Scatter plot shows correlation of SIX2-_rep1 ChIP-seq reads and SIX2-_rep2 within +/- 500 bp around SIX2-rep1 ChIP-seq peaks (left) and vice versa (right).
Figure S2: Human and mouse SIX2/Six2 ChIP antibodies are comparable. (A) Scatter plot shows comparison of Six2 ChIP-Seq data generated using the mouse Six2 ChIP antibody (Six2-ms-ab) or SIX2 human ChIP antibody (SIX2-hu-ab). Average enrichment fold (ChIP-Seq score) of replicate ChIP-Seq reads count within +/-500 bp of the indicated peaks were plotted. We identified ‘shared peaks’ and ‘Six2-ms-ab only peaks’ based on ChIP-Seq score, with metrics specified in the Methods. (B) Weblogo shows the most significantly enriched motif identified from +/-50 bp window of top 1000 shared peaks; no motif was identified from Six2-ms-ab only peaks. Percentages indicate frequency of the motif appearing in the indicated peaks. P-values indicate statistical significance of the frequency. (C) Barplots show distribution of distances of peaks to the nearest gene. (D) Barplots show enrichment of Gene Ontology (GO) terms for the shared peaks and Six2-ms-ab only peaks. The analysis was performed using the GREAT on-line tool (McLean et al., 2010).
Figure S3: Human and mouse SIX2 binding around EYA1 suggests EYA1 is key target of SIX regulation. (A) Genomic view of human and mouse SIX2 binding near EYA1/Eya1 gene loci. Human tracks represent enrichment fold for SIX2 ChIP-seq, H3K27ac ChIP-seq and RNA-seq from ITGA8+ nephron progenitors, mouse tracks represent Six2 and H3K27ac ChIP-seq from E16.5 whole kidneys and H3K27me3 ChIP-seq and RNA-seq from Cited1+ nephron progenitors. 'Cons' track represents the Phastcon vertebrate conservation score. Highlighted regions are conserved from mouse to human. Mouse Six2 peak annotated with * was previously tested in a G0 transgenic assay and recapitulated Eya1 expression (Park et al., 2012).
Figure S4: ITGA8 in the human kidney, identification of SIX1 and SIX2 specific antibodies, and comparison of SIX1/Six1 and SIX2/Six2 protein. (A) ITGA8 staining in a 16 week human fetal kidney before and after limited enzymatic digest for FACS showing the removal of the ITGA8+ nephron progenitors. Bottom panels show a cytospin sample cells collected from the ITGA8 positive FACS population or the negative population. (B) NIH3T3 cells were transfected with constructs to induce expression of SIX1 or SIX2. GFP was utilized as a control. Cells were stained with a SIX2 antibody raised against the full-length protein previously utilized for Six2 ChIP-seq (Park et al., 2012; top panel), a SIX2 antibody raised against a C-terminal epitope (aa264-277) (middle panel), and a SIX1 antibody raised against residues surrounding Pro249 (bottom panel). (C) Alignment of SIX1/Six1 and SIX2/Six2.
Figure S5: Additional expression patterns observed from the injection of SIX1 Enhancer
1. G0 transgenic analysis was performed at E15.5 and urogenital systems stained for LacZ reporter expression. Varied expression patterns are observed, with no two recapitulating the same pattern.
**Figure S6: Analysis of SIX1 ChIP-Seq replicates and SIX1/SIX2 ChIP-seq correlation.** (A) Venn diagram shows overlap (100 bp gap) between two replicates of SIX1 ChIP-seq peaks from 16 (rep1) and 17 week (rep2) human fetal kidneys. (B) Scatter plots show high correlation of the two SIX1 tag counts within SIX1-rep1 or SIX1-rep2 binding regions. (C) Histogram shows distribution of motif-peak distances of the SIX1 peaks from the two replicates. The motif is centered in both datasets, indicating direct association of SIX1 with the motif. (D) Scatter plot shows a positive correlation of SIX1 and SIX2 reads within +/- 500 bp of SIX1 peaks. (E) Boxplot shows distribution of the ChIP-seq binding score for peaks unique to SIX1 (left) and SIX1 peaks overlapping with SIX2 and peaks (right). The size of the boxes are proportional to the square roots of the numbers of observations. Medians of each group are annotated on the left side of the boxes and indicate a higher binding score for those peaks that are shared between SIX1 and SIX2. (F) Motif discovery in SIX1-only, SIX2-only and SIX1/SIX2 co-binding sites.
Table S1: Details of samples, antibodies, and basic analyses for ChIP-seq and RNA-seq experiments.

<table>
<thead>
<tr>
<th>Sample name</th>
<th>Sample type</th>
<th>Species</th>
<th>Sample stage/tissue</th>
<th>ChIP Ab</th>
<th>Note</th>
<th>Mapped reads</th>
<th>Threshold</th>
<th>Peaks</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIK2-repl1</td>
<td>ChIP-seq</td>
<td>Human</td>
<td>17 week kidney cortex</td>
<td>SIK2 (Mybiosource, MBS610128)</td>
<td>Ab specific to SIK2</td>
<td>29120274</td>
<td>50</td>
<td>6275</td>
<td>0</td>
</tr>
<tr>
<td>SIK2-repl2</td>
<td>ChIP-seq</td>
<td>Human</td>
<td>17 week kidney cortex</td>
<td>SIK2 (Mybiosource, MBS610128)</td>
<td>Ab specific to SIK2</td>
<td>25730785</td>
<td>10</td>
<td>54068</td>
<td>0.02</td>
</tr>
<tr>
<td>SIK1-repl1</td>
<td>ChIP-seq</td>
<td>Human</td>
<td>16 week kidney cortex</td>
<td>SIK1 (Cell Signaling, 12891S)</td>
<td>Ab specific to SIK1</td>
<td>24821763</td>
<td>10</td>
<td>4433</td>
<td>0.11</td>
</tr>
<tr>
<td>SIK1-repl2</td>
<td>ChIP-seq</td>
<td>Human</td>
<td>17 week kidney cortex</td>
<td>SIK1 (Cell Signaling, 12891S)</td>
<td>Ab specific to SIK1</td>
<td>2659792</td>
<td>10</td>
<td>6535</td>
<td>0.16</td>
</tr>
<tr>
<td>H3K27ac</td>
<td>ChIP-seq</td>
<td>Human</td>
<td>17 week kidney cortex</td>
<td>H3K27ac (Abcam, ab4729)</td>
<td></td>
<td>48190122</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SIK2</td>
<td>ChIP-seq</td>
<td>Mouse</td>
<td>E16.5 whole kidney</td>
<td>SIK2 (Proteintech, 11562-1-AP)</td>
<td>Ab recognizes SIK1 and SIK2</td>
<td>30936440</td>
<td>60</td>
<td>12145</td>
<td>0.0003</td>
</tr>
<tr>
<td>H3K27ac</td>
<td>ChIP-seq</td>
<td>Mouse</td>
<td>E16.5 whole kidney</td>
<td>H3K27ac (Abcam, ab4729)</td>
<td></td>
<td>21399295</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>ChIP-seq</td>
<td>Mouse</td>
<td>E16.5 Cited1+ cells</td>
<td>H3K27me3 (Abcam, ab6002)</td>
<td></td>
<td>213833317</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Human-input1</td>
<td>ChIP-seq</td>
<td>Human</td>
<td>17 week kidney cortex</td>
<td>NA</td>
<td></td>
<td>44091209</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Human-input2</td>
<td>ChIP-seq</td>
<td>Human</td>
<td>16 week kidney cortex</td>
<td>NA</td>
<td></td>
<td>40720333</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Mouse-input</td>
<td>ChIP-seq</td>
<td>Mouse</td>
<td>E16.5 whole kidney</td>
<td>NA</td>
<td></td>
<td>19679957</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Progenitor RNA</td>
<td>RNA-seq</td>
<td>Human</td>
<td>17 week iTGAB+</td>
<td>NA</td>
<td></td>
<td>11718597</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Progenitor RNA</td>
<td>RNA-seq</td>
<td>Mouse</td>
<td>E16.5 Cited1+ cells</td>
<td>NA</td>
<td></td>
<td>57908653</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table S2

Click here to Download Table S2

Table S3

Click here to Download Table S3

Table S4

Click here to Download Table S4