

## Genome-wide identification of CBX2 targets: insights in the human sex development network

Wassim Eid<sup>1</sup>, Lennart Opitz<sup>2</sup>, Anna Biason-Lauber<sup>1</sup>

<sup>1</sup>Division of Endocrinology, Department of Medicine, University of Fribourg, 1700 Fribourg, Switzerland;

<sup>2</sup> Functional Genomics Center Zurich (FGCZ), University of Zurich/ETHZ, 8057 Zurich, Switzerland.

CBX2 (Chromobox homolog 2) is a chromatin modifier that plays an important role in sexual development and its disorders (disorders of sex development, DSD), yet the exact rank and function of human CBX2 in this pathway remains unclear. Here, we performed large-scale mapping and analysis of *in vivo* target loci of the protein CBX2 in Sertoli-like NT-2D1 cells, using the DNA adenine methyltransferase (DamID) technique. We identified close to 1600 direct targets for CBX2. Intriguingly, validation of selected candidate genes using qRT-PCR in cells overexpressing CBX2 or in which CBX2 has been knocked down indicated that several CBX2-responsive genes encode proteins that are involved in DSD. We further validated these effects on the candidate genes using a mutated CBX2 causing DSD in human patient. Overall, our findings suggest that CBX2 role in the sex development cascade is to stimulate the male pathway and concurrently inhibit the female pathway. These data provide fundamental insights into potential etiology of DSD.

The process of sexual differentiation is central for reproduction in almost all metazoan, and therefore, for maintenance of practically all multicellular organisms. The critical step in this process is the determination of the bipotential embryonic gonad into either ovaries or testes. Defects in any of the genes involved in either the testicular or ovarian development can result in disorders of sex development (DSD) (1, 2). Although there have been considerable advances in our understanding of the genetic factors involved in sexual differentiation, the fact remains that in most DSD the underlying genetic cause is unknown, in fact it has been estimated that a molecular diagnosis is made in only about 20% of DSD (3, 4). This highlights the substantial gap in knowledge that needs to be addressed.

CBX2 (chromobox homolog 2) was recently added as an additional player in human sex development process (5). CBX2 and its mouse homologue M33 are members of the Polycomb (PcG) family of proteins. PcG proteins are highly conserved transcriptional regulators that form large multiprotein complexes which exert their function through the modulation of higher order chromatin struc-

ture (6, 7). They are required for stable maintenance of gene expression programs through developmental and cell differentiation processes. Although best known for their role in controlling the expression of *Hox* genes in *Drosophila* during development, numerous PcG members have been now implicated in the control of various cellular processes, such as chromosome X-inactivation, cell fate decisions, tumorigenesis and sexual differentiation (8–12). Human CBX2 exists in two isoforms, a 532-amino acid long isoform (NP\_005180), called CBX2.1 (hereafter referred to as CBX2), and a second shorter 211-amino acid isoform (NP\_116036), named CBX2.2. The importance of M33/CBX2 for sexual development is highlighted by the fact that M33-deficient mice have male-to-female sex reversal (13). Similarly, in humans we made the intriguing discovery of a CBX2 loss-of-function double heterozygote mutation state in a 46, XY girl with ovarian tissue at histology, normal uterus and external genitalia, accidentally diagnosed because of discrepancy between prenatal karyotypes and phenotype at birth. Functional studies demonstrated that the mutated CBX2 does not properly bind to and does not adequately regulate the

expression of target genes essential for sex development such as *SF1/NR5A1* (5), which was in agreement with other data (14).

Although these data elucidated that CBX2 is involved in the molecular pathogenesis of DSD and that it lies upstream of SRY in the sex determination cascade, little is known about the exact rank of CBX2 and its role in this cascade. To gain insight into CBX2 function, we have employed the DNA adenine methyltransferase identification (DamID) coupled with next-generation sequencing (NGS) to analyze the human CBX2 localization across the genome and identify its direct targets.

In this study, we identified close to 1600 direct targets of CBX2 in human Sertoli-like cells. Furthermore, we provide evidence that the expression of some of the identified targets is CBX2-dependent, thereby identifying potential genes that could be involved in DSD in humans and setting the ground for future studies.

## Materials and Methods

### Cell culture

HEK293T (ATCC CRL-11 268) and NT2-D1 (ATCC CRL-1973) cells were obtained from the American Type Culture Collection and cultured at 37°C, 5% and 10% CO<sub>2</sub> respectively, in Dulbecco's Modified Eagle Medium containing 10% fetal bovine serum and supplemented with penicillin-streptomycin.

### siRNA

siRNA duplexes were purchased from Microsynth (Switzerland), siRNA sequences are as follows: (si Scrambled) (CGUACGCGGAAUACUUCGATT) (15) CBX2-145(GGCUGGUCCUCCAAACAUATT) and CBX2-411(GGAUGACAGUGAGUUAGAUTT) (5). siRNA duplexes were introduced into cells using Lipofectamine RNAiMAX (Invitrogen) in two consecutive rounds at a final concentration of 40 nM. Experiments were typically performed 48h post the first transfection. For targets that did not show a significant change in expression at 48h, experiments were repeated at 72h.

### Immunofluorescence

Cells grown on cover slips were pre-extracted for 5 minutes on ice using 25 mM HEPES pH 7.4, 50 mM NaCl, 1 mM EDTA, 3 mM MgCl<sub>2</sub>, 300 mM sucrose and 0.5% Triton X-100 before fixation in 4% formaldehyde (w/v) in PBS for 15 minutes at room temperature (RT). Cover slips were incubated overnight at 4°C with primary antibodies and Alexa-conjugated secondary antibodies (check SI for detailed antibodies information) for 1h at RT. The cover slips were mounted with Vectrashield (Vector Laboratories) containing DAPI. Images were acquired using a Nikon eclipse Ni-E microscope system.

### DamID

DamID was performed using the lentiviral transduction protocol as previously described (16), using pLgw-EcoDam-V5-

CBX2 (EcoDam-CBX2) or pLgw-V5-EcoDam (Dam-only). Briefly, EcoDam-CBX2 or Dam-only were transfected into HEK293T cells together with the packaging plasmids (pRSV-Rev, pCMV-dR8.2, pMD2.G) using Metafectene (Biontex Laboratories), 48h later the supernatant containing the lentivirus was harvested and added to NT-2D1 cells with polybrene 1 mg/ml (Santa Cruz Biotech), two consecutive rounds of transduction were performed. Two days after the first infection, genomic DNA was isolated using DNeasy blood and tissue kit (Qiagen). The isolated gDNA was subject to DpnI digestion, ligation of DamID adaptors and DpnII digestion. Samples without addition of DpnI or ligation adaptors served as negative controls. DpnII-digested fragments were used as templates to amplify methylated genomic fragments.

### Next-Generation sequencing (NGS)

Library preparation was performed using standard procedures, see supporting information (SI) for a detailed protocol. The sequence alignment to the human genome (hg19) was performed with Bowtie2 using the options '-sensitive' and '-non-deterministic'. To identify peaks from the data we used MACS (version 1.4) with default options. The subsequent analysis including the annotation and visualization of peaks was done with the Bioconductor-packages ChIPpeakAnno and Gviz from R.

### Quantitative real time PCR

Total RNA was extracted using the RNeasy Mini Kit (Qiagen) and reverse-transcribed using the Omniscript reverse-transcription system (Qiagen). Experiments were performed with an ABI StepOnePlus Real-Time PCR system, PCR products were quantified fluorometrically using the KAPA SYBR FAST master mix (KAPA Biosystems). The reference mRNA cyclophilin was used for normalizing the data (primers sequence available upon request). All samples were run in triplicates. Unpaired *t* test was performed using GraphPad Prism version 5.04 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com.

## Results

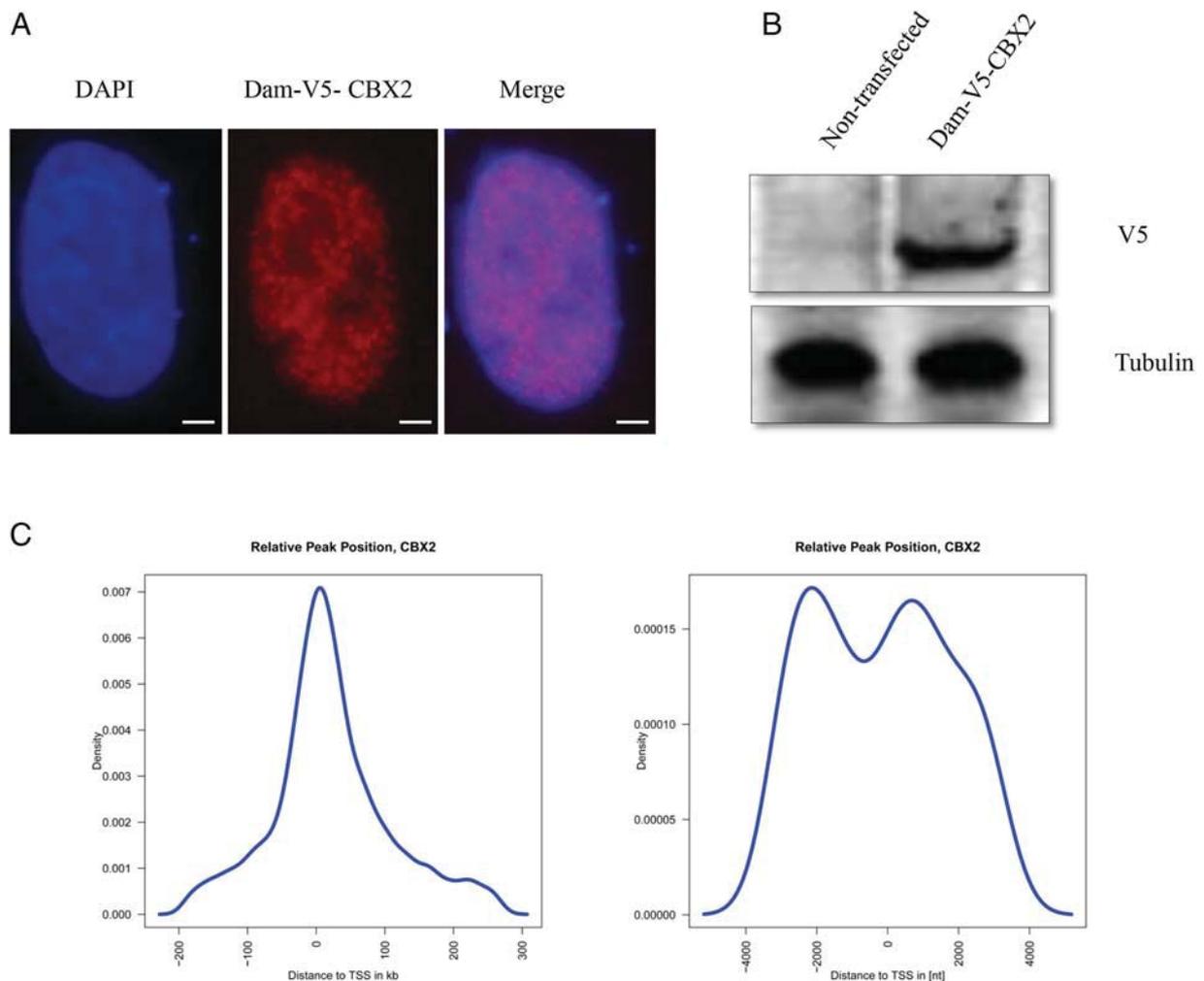
### Mapping of CBX2 target loci

To gain insights into the function of CBX2 and to identify its targets, we applied the DamID-seq method that couples next-generation sequencing (NGS) to DamID assay. DamID is a methodology independent of antibodies, fixation, chromatin shearing and other technically challenging features of Chromatin Immuno Precipitation (ChIP) a technique commonly used to explore DNA/protein interactions. It has proven to be a powerful tool for the mapping of genomic binding sites of chromatin proteins (16). First, we tethered human CBX2 to *E.coli* DNA adenine methyltransferase (Dam) and expressed the fusion protein in human testis-derived NT-2D1 cells. In eukaryotic cells that lack endogenous Dam proteins, the tethered Dam is capable of methylating adenines in GATC sequences that are in close spatial proximity (17).

We confirmed by Western blot and immunofluorescence that the exogenously introduced CBX2-Dam fusion protein was expressed and properly localized to the nucleus (Figure 1A and B). Next, we transduced NT-2D1 cells with the lentivirus expressing Dam-CBX2, purified the gDNA from transduced cells, specifically amplified the methylated DNA fragments and sequenced them. Cells expressing free Dam protein served as a control to correct for the unspecific DNA accessibility by Dam alone and for sequencing biases. We prepared two replicates for each fusion protein or free Dam and sequenced the DNA libraries separately. All pairs of replicates had highly correlated read densities along the genome (Pearson correlation coefficients  $> 0.90$ ,  $p$ -value  $< 2e-16$ ) and thus were combined for the follow-up analyses. In total, we obtained 85 million and 17 million reads from NT-2D1

expressing Dam-CBX2 and free Dam, respectively. After peak-calling we identified 1595 direct targets of CBX2 with a  $p$ -value  $< 1e-5$  (Supplemental Table 1).

In order to elucidate the detailed distribution of CBX2 binding we analyzed the identified putative target regions of CBX2. The distance to the closest Transcription Start Site (TSS) showed that most of all the called peaks is nearby a TSS (Figure 1C). A detailed view on the 5Kb up- and downstream region around the TSS showed depletion in the actual start site while the largest peak is located in the promoter region of the putative target genes (Figure 1C). Analysis of the putative CBX2 binding sites (Figure 2A) confirms that most peaks are upstream to the closest TSS. The fraction of peaks found in the promoter region defined by 5Kb upstream is rather small ( $< 3\%$ ) (Figure



**Figure 1.** The DamID platform for location analysis of human CBX2. (A) Dam-CBX2 localizes to the nucleus. NT2-D1 cells were transfected with the pLgw-EcoDam-V5-CBX2 (Dam-CBX2), 48 hours later cells were detergent-extracted and fixed. Dam-CBX2 was visualized by indirect immunofluorescence using a V5 antibody, nuclei were visualized by DAPI. Bar 5  $\mu$ m. (B) Immunoblot analysis of Dam-CBX2 expression in NT2-D1 cells. Whole cell extracts of cells transfected with same construct as in A were analyzed by immunoblotting for the expression of Dam-CBX2 using a monoclonal antibody against V5; tubulin is used as a loading control. (C) Identified enriched regions localize relatively close to transcription start sites. In a 5Kb window around the TSS are two peaks in the distribution. The highest peak reflects binding events in the promoter region. Considering all identified putative binding regions there is a general peak close to a TSS.

2B). In contrast many peaks (>65%) are located in intergenic regions.

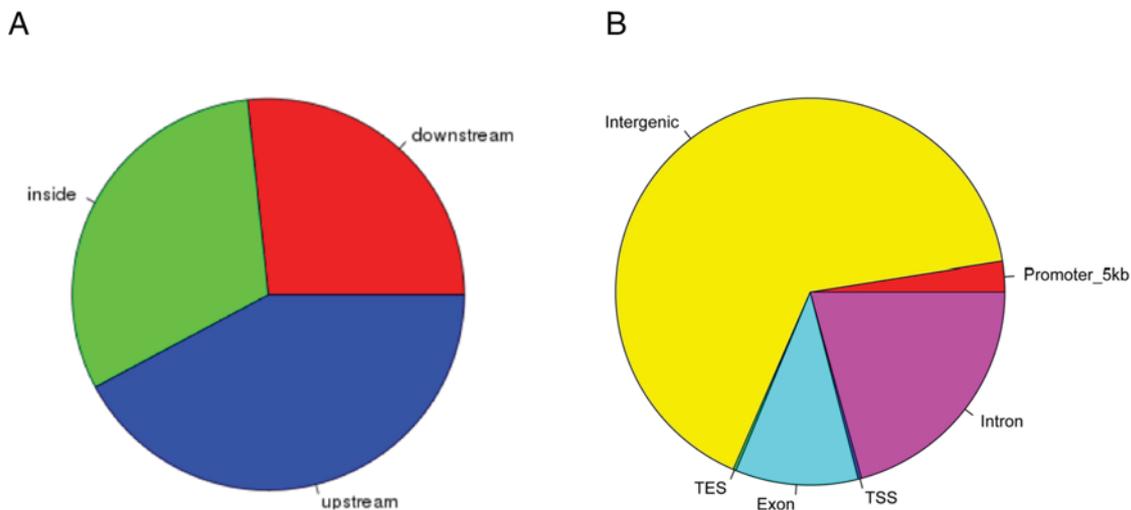
### Validation of CBX2 targets

To independently test the DamID-seq results, we selected a subset of CBX2 targets identified by DamID-seq for further investigation and validation. Target selection was based on their reported putative involvement in sexual differentiation, their biological role in human disease and animal models or their specific expression in tissues involved in sex development (gonads, sex organs, hypothalamus and pituitary) (Table 1). Applying these selection criteria we short-listed nineteen genes. In basal (control) conditions, expression analysis showed NT2-D1 cells endogenously expressed fifteen of these genes (*SOX9*, *MAMLD1*, *EXO1*, *SOX3*, *TEX10*, *TBX2*, *PBX1*, *TSNAX*, *FGF2*, *FGFR2*, *ATRX*, *WTAP*, *MTM1*, *TSPYL4*, *FZD1*) while four genes namely *KISS1*, *ZP4*, *PASD1* and *IGF1* were not significantly expressed in this particular cell line. (Figure 4 and not shown). To study the influence of CBX2 on the expression of the selected genes we measured their expression levels in NT2-D1 cells under overexpression and upon selectively knocking down endogenous CBX2 using specific siRNAs (5), respectively. The expression level of CBX2 upon its exogenous overexpression or its knocking down was confirmed at the protein and mRNA levels by Western blotting and qRT-PCR analyses respectively (Supplemental Figure 1).

As a prelude for validating the candidate target genes, we first checked the effect of CBX2 on genes known to be regulated by it such as *SOX9*, *SF1* and *SRY* (5, 11). *SOX9* was among the identified targets in our DamID-seq. Compared to control transfection (empty vector), overexpression of CBX2 enhanced expression of *SOX9*, *SF1*/

*NR5A1* and *SRY* by 4, 3 and 3-fold respectively, while knocking down endogenous CBX2 resulted in down regulating the expression of the same genes to 0.6, 0.5 and 0.6-fold respectively (Supplemental Figure 2), consistent with previous observations (5).

Next, we examined the effect of CBX2 on the selected candidate genes. The wild-type (WT) CBX2 significantly enhanced the expression of endogenous *EXO1*, *SOX3*, *MAMLD1*, *TBX2*, *TEX10*, *FGF2*, *WTAP*, *TSPYL4* and *MTM1* (by 2.9, 2.6, 2.6, 2.3, 2.2, 2.4, 1.6, 2.2, and 1.6-fold respectively), but decreased the expression of *PBX1* and *FZD1* (to 0.4, 0.5 and 0.7-fold respectively) (Figure 4 and 5). The expression of *TSNAX*, *FGFR2* and *ATRX* seemed not to be affected by CBX2 (Supplemental Figure 3). To further explore the specificity of these observations, we knocked down endogenous CBX2 expression by siRNA oligos. Expression analysis showed that down regulation of CBX2 significantly reduced the expression of *EXO1*, *SOX3*, *MAMLD1*, *TBX2*, *TEX10* and *TSPYL4* (to 0.7, 0.6, 0.6, 0.6, 0.7 and 0.8-fold respectively), while enhancing the expression of *PBX1* and *FZD1* (by 2.3, 1.9 and 1.4-fold respectively) compared to control (scrambled siRNA), further supporting the CBX2-dependent expression of these genes. The lack of significant effects on the expression of *MTM1*, *WTAP* and *FGF2* after down regulating CBX2 is unclear but it might be attributed to redundant pathways controlling the expression of these particular genes. Expression analysis for *MTM1*, *WTAP* and *FGF2* were re-examined at different time points of CBX2 downregulation, no significant change in their expression levels were observed (data not shown).



**Figure 2.** Genomic binding features of CBX2. (A) Localization of identified enriched binding regions in relation to described gene models. (B) The enriched regions are representing different genomic features. Based on the overlap an identified region can be assigned to more than one feature type. The promoter region is defined as 5Kb upstream of a TSS.

**Table 1.** CBX2 targets selected for expression profiling

Gene Symbol	Gene name	Function	Link to sex development	Fold change upon CBX2 overexpression (EV = 1)	Fold change upon CBX2 knockdown (siScr = 1)
SOX9	SRY (sex determining region Y)-box 9	Downstream target of SRY	Mutation lead to 46, XY DSD (12)	↑ 4.0	↓ 0.6
MAMLD1	Mastermind-like domain-containing protein 1	Transactivates the HES3 promoter independently of NOTCH proteins	Mutation leads to hypospadias (40)	↑ 2.6	↓ 0.6
TEX10	Testis-expressed sequence 10 protein	Functions as a component of the Five Friends of Methylated CHTOP (5FMC) complex	Downstream binding target of SRY (20)	↑ 2.2	↓ 0.7
EXO1	Exonuclease 1	Essential for male and female germ cells meiosis	Loss leads to male and female sterility (23)	↑ 2.9	↓ 0.7
SOX3	Transcription factor SOX-3	Required during the formation of the hypothalamo-pituitary axis	Mutation leads to 46,XX DSD (41)	↑ 2.6	↓ 0.6
TBX2	T-box transcription factor TBX2	Involved in the transcriptional regulation of genes required for mesoderm differentiation	Might be involved in development of the reproductive system (21)	↑ 2.3	↓ 0.6
PBX1	Pre-B-cell leukemia transcription factor 1	Transcriptional activator	Loss leads to reduced urogenital ridge outgrowth and absence of Müllerian ducts (35)	↓ 0.5	↑ 1.9
FGFR2	Fibroblast growth factor receptor 2	Tyrosine-protein kinase that acts as cell-surface receptor for fibroblast growth factors	Loss leads to partial XY sex reversal (42)	n.s	n.s
FGF2	Fibroblast growth factor 2	Potent mitogen and neurotropic factor for brain cells in vitro	Ligand of FGFR2	↑ 2.4	n.s
WTAP	WT1-associated protein	Impairs WT1 DNA-binding ability and inhibits expression of WT1 target genes.	Mutation might be involved in hypospadias (43)	↑ 1.6	n.s
ATRX	Transcriptional regulator ATRX	Global transcriptional regulator	XY patients with deletions or mutations have varying degrees of sex reversal (44)	n.s	n.s
MTM1	Myotubularin	Lipid phosphatase	Deletion leads to myotubular myopathy and abnormal genital development (25)	↑ 1.6	n.s
TSPYL4	Testis-specific Y-encoded-like protein 4	Belongs to the nucleosome assembly protein (NAP) family	TSPYL mutation leads to sudden infant death and dysgenesis of the testes (22)	↑ 2.2	↓ 0.8
FZD1	Frizzled-1	Receptor for Wntproteins	Loss leads to sterility in females (30)	↓ 0.7	↑ 1.4
TSNAX	Translin-associated protein X	Acts in combination with TSN as an endonuclease involved in the activation of the RNA-induced silencing complex (RISC)	Null male mice have abnormal seminiferous tubules and reduced sperm count. Null female mice are subfertile (45)	n.s	n.s

The two right most columns indicates the behaviour of the candidate target gene in response to CBX2 as up (↑)- or down (↓)- regulated. Change in expression are represented as relative to control = 1. (EV) empty vector, (siScr) si scrambled, (n.s) non-significant.

In whole transcriptome studies in NT2-D1 cells we observed that CBX2 negatively regulates *FOXL2* expression (unpublished data). The validation of this results showed a slight but significant reduction in the expression of *FOXL2* in response to overexpressed CBX2 and reciprocally, a significant up regulation of its expression when CBX2 was knocked down (Figure 5).

To study the pathophysiological consequences of CBX2-defects, we took advantage of the mutations we described as being the molecular basis of 46, XY DSD in human patients (5) and repeated the experiments by introducing the mutated forms of CBX2 harboring a P98L and R443P mutations into NT2-D1. Immunofluorescence imaging showed that both mutated forms localize to the nucleus and qRT-PCR analysis suggested that there are no significant differences in expression between WT and mutated CBX2 (Supplemental Figure 3). In contrast to WT CBX2, the mutated CBX2 had no effect on the expression of all the selected targets and behaved essentially as the control. (Figure 4 and 5).

## Discussion

Chromatin regulators are critical integration points wherein biological signals are converted into changes in

gene expression. These protein machineries are essential to normal cell function, development and differentiation (18). Critical for understanding the biology of these regulators is the determination of their binding sites to deduce their actions within the genome and eventually understand how the subsequent cellular responses occur. This is of particular importance in case of disorders of sex development (DSD), where albeit the advance in understanding the genetic basis of human sexual determination and differentiation, in most subjects with DSD, the underlying genetic cause is unknown. Whereas pathogenic mutations have been identified in DSD patients, the phenotype can be highly variable, even within families, suggesting that other factors, including genetic variants are influencing the expression of the phenotype (3). Unexplained DSD cases may be due to mutations in novel gonad-determining genes, or genomic alterations affecting regulatory regions that lead to atypical gene expression. The identification of new genes involved in sex determination is therefore a goal of the utmost importance to the understanding of DSD and will allow formore accurate diagnosis and clinical management of these complex conditions (2).

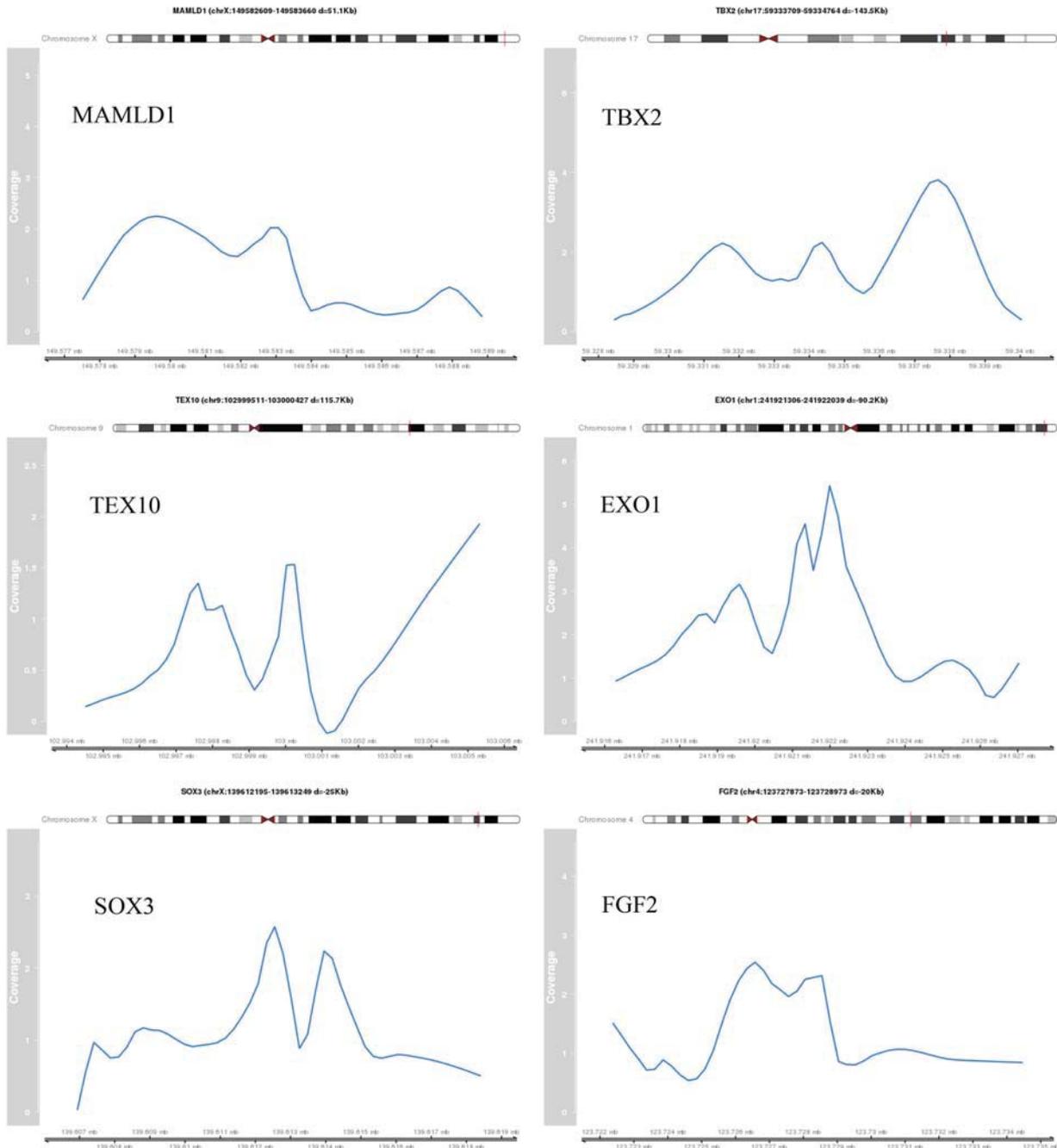
To this goal, we chose to generate the first genome-wide binding data for human CBX2, a chromatin modu-

lator that was shown to be involved in sex development in mouse (11) and humans (12). To explore its precise position in this developmental cascade we employed the DamID assay coupled with next-generation sequencing (DamID-seq) that enabled us to identify 1595 direct targets for CBX2 in the human Sertoli-like cells NT-2D1. To further refine our identification of targets, and to place these genes in the context of the sexual determination pathway, we performed expression profiling experiments on a subset of genes in cells overexpressing CBX2 or in

cells in which endogenous CBX2 has been selectively knocked down.

Among the factors known to play a role in male sex development, we found *SOX9* to be among our identified targets of CBX2. *SOX9* is the first identified target of *SRY* and its expression is essential for testis development (19). *SOX9* expression has been linked to CBX2 in mouse studies (11), as well as in human cells (5), yet it remained unclear if CBX2's influence on *SOX9* is direct or indirect. By identifying *SOX9* sequences as targets for CBX2 bind-

http://doc.rero.ch

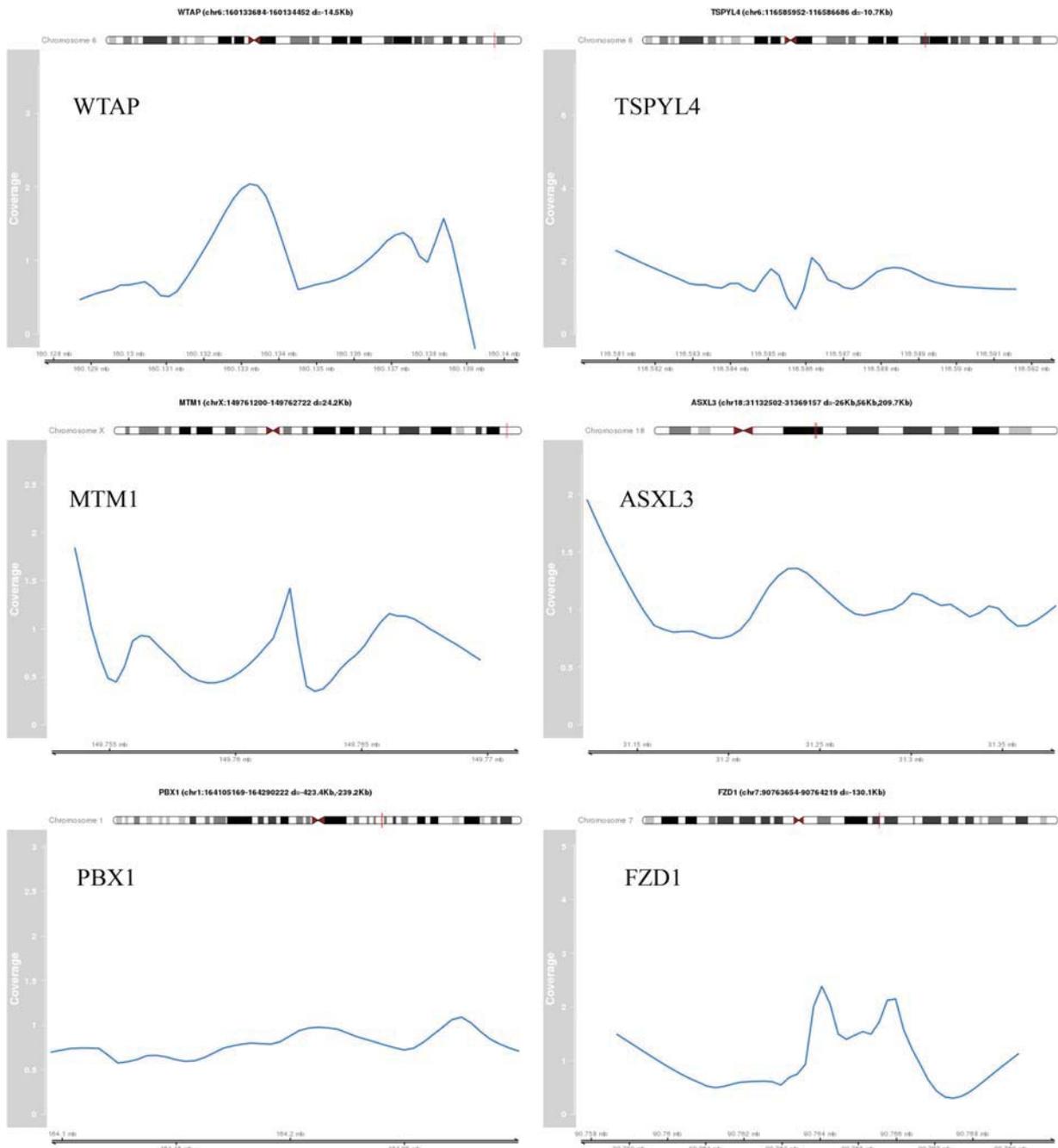


**Figure 3a.** Identified peak regions of validated target genes. The coverage profile of the peak region normalized by the control sample is shown for selected validated target genes.

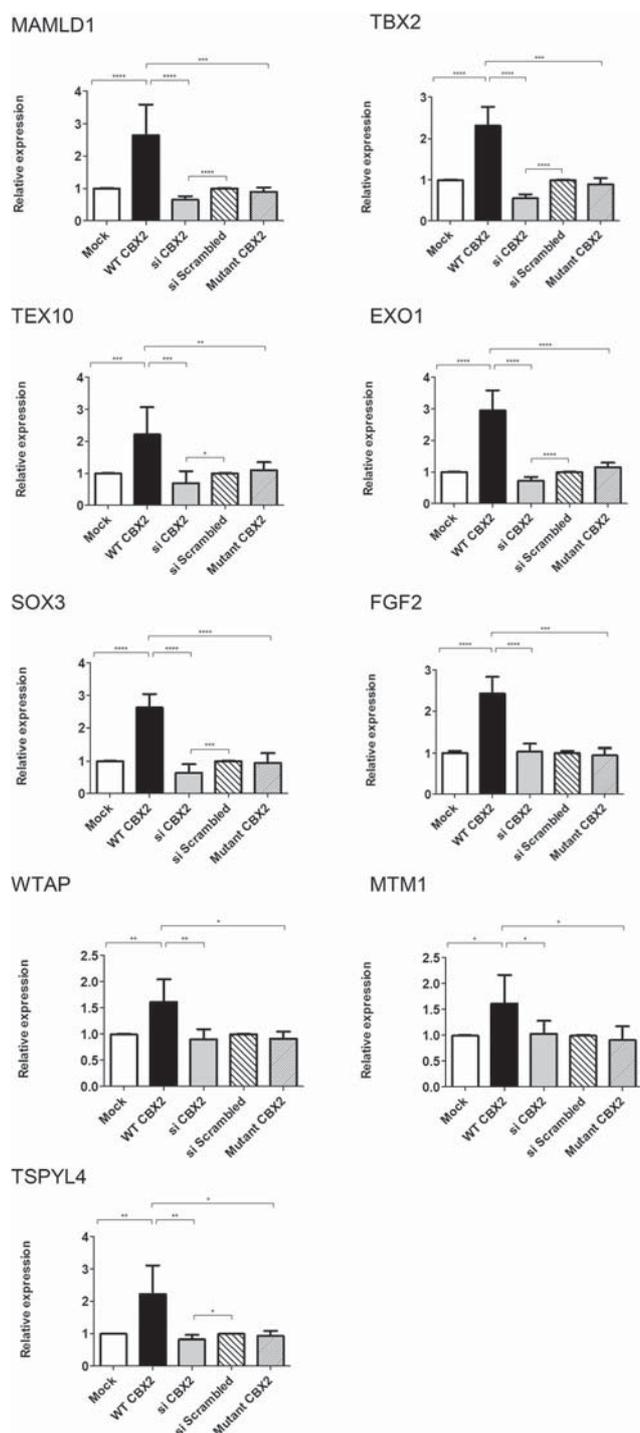
ing, our finding establishes the first direct link between CBX2 and *SOX9*. Accordingly, the expression of *SOX9* were increased by approximately 4-fold upon overexpressing CBX2, while decreased to 0.6-fold upon knocking down endogenous CBX2, which is agreement with previous data (5).

Numerous genes identified by our DamID-seq were linked to DSD in humans: *MAMLD1*, *SOX3*, *FGFR2*, *ATRX* (2), but their CBX2-dependent regulation is a novel finding. In addition, further analysis of the genomic landscape surrounding CBX2 revealed a number of addi-

tional gene candidates such as *TEX10*, *EXO1*, *TBX2*, *TSPYL4*, *WTAP* and *MTM1*, suggesting an additional role of CBX2 in sexual development. For example, *TEX10* has been found to be a downstream binding target of *SRY* (20). *TBX2* is involved in the transcriptional regulation of genes required for mesoderm differentiation, is overexpressed in reproductive system cancers and has been suggested to play a role in the development of the mammalian reproductive system (21), while *TSPYL4* was reported to be potentially linked to testicular dysgenesis (22). *EXO1* is essential for male and female meiosis with



**Figure 3b.** The genomic coordinates are indicating the peak region. The distance (d) to the TSS of the target gene is provided.



**Figure 4.** Genes up-regulated in a CBX2-dependent manner. qRT-PCR quantification of mRNA levels for putative CBX2 targets that are up-regulated by CBX2. NT2-D1 cells transfected with empty vector (Mock, white) or wild-type (WT) CBX2 (black) or mutant CBX2 (narrow lines) or scrambled siRNA (wide lines) or siRNA against CBX2 (gray). Relative expression levels of target genes were determined by qRT-PCR after normalization to cyclophilin as an endogenous control. Following CBX2 overexpression (black) the relative expression of *MAMLD1*, *TEX10*, *EXO1*, *SOX3*, *TBX2*, *FGF2*, *WTAP*, *MTM1*, *TSPYL4* was up-regulated compared to cells transfected with empty vector (white). CBX2 knockdown using siRNA (gray) reduced the relative expression of *MAMLD1*, *TEX10*, *EXO1*, *SOX3*, *TBX2*, *TSPYL4*, while not significantly affecting the expression of *FGF2*, *WTAP*, *MTM1* compared to

*Exo1* *-/-* mice being sterile (23). WTAP functions by impairing Wilms tumor protein-1 (WT1) DNA-binding and inhibits expression of WT1 target genes (24). *MTM1* deletion has been associated with myotubular myopathy and abnormal genital development (25).

We were unable to detect any significant changes in expression of *FGFR2*, *WTAP* and *MTM1* after knock-down of CBX2, despite the fact that their expression levels were significantly increased upon overexpressing CBX2. According to these results CBX2 may only have an secondary role in the regulation of these genes. However, it should be noted that the siRNA-mediated knockdown of CBX2, although substantial, may have been insufficient or not long enough to cause detectable alteration in gene expression.

Few of our identified genes (*TSNAX*, *FGFR2* and *ATRX*), showed no significant change in expression levels in response to CBX2 overexpression or knockdown. For *ATRX*, a previous study reported a similar observation where the expression of the murine *Atrx* was not significantly affected in M33/CBX2 knock-out mice (11). This suggests that additional factors, perhaps coactivators and chromatin state, are involved in their transcriptional regulation.

In the same line of reasoning, we observed that among the identified CBX2 targets are other transcription factors (*TBX2*, and *ATRX*). Accordingly, previous reports suggested that CBX2 regulates other genes, particularly *SRY* in an indirect fashion through the regulation of other genes encoding transcription factors (11). We and others (11) could not find a direct interaction between CBX2 and *SRY* that could explain the enhanced expression of *SRY* in response to CBX2. It is intriguing to speculate the association of any of the identified transcription factors in the CBX2-dependent *SRY* regulation. Supporting this idea, two recent studies in rats lacking the Y chromosome and the *SRY* gene, suggested that additional copies of CBX2 in males might be involved in a novel sex-determining mechanism in species that lack *SRY* (26, 27). This observation underlines the central role of CBX2 in male sexual development as a factor independent from *SRY*, although additional work is still needed in this perspective.

We also identified CBX2 binding near a considerable

**Legend to Figure 3b Continued. . .**

scrambled siRNA (wide lines). Introducing the mutated forms (P98L and R443P) of CBX2 (narrow lines) in equimolar amounts into the cells showed no effect on the expression of all the target genes. The data in all graphs are the average of three independent experiments, error bars represent standard deviation from the mean and values are expressed as relative to control = 1. (\*\*\*\*)  $P < .0001$ ; (\*\*\*)  $P < .001$ ; (\*\*)  $P < .01$ ; (\*)  $P < .05$ .

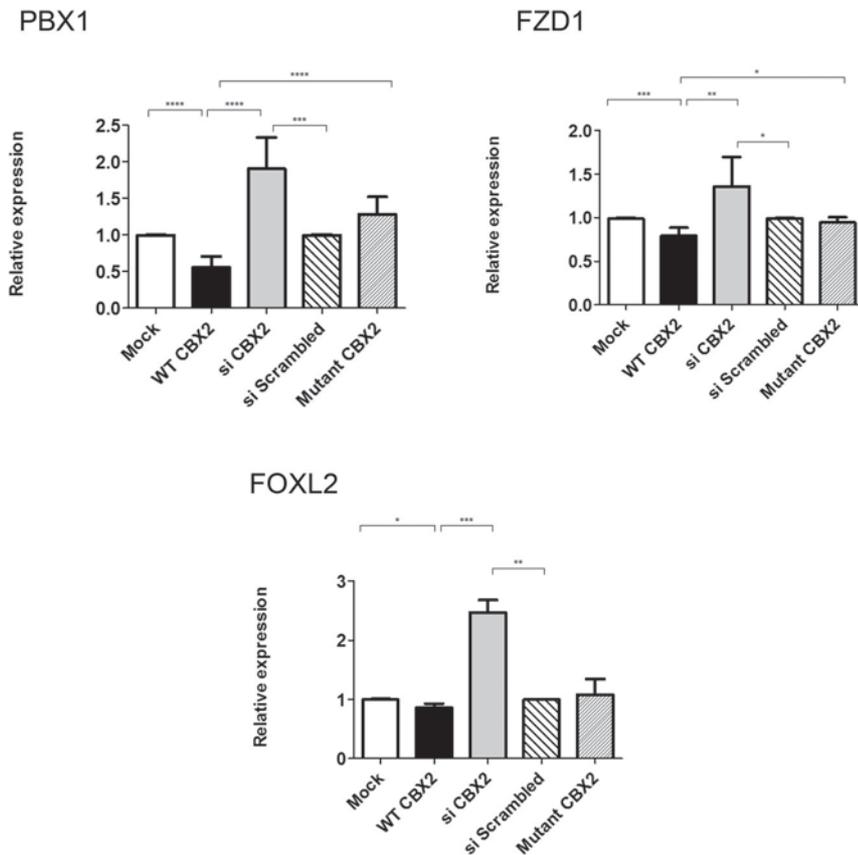
number of genes encoding microRNAs, implying a possible still undiscovered role of microRNAs in human sex development, similarly to what has been shown for microRNAs and testicular proteome in mice (28), as well as in gonadal sex differentiation in chicken embryos (29). Further characterization of the impact of CBX2 activation on microRNAs and their potential involvement in sexual development in humans is essential.

Among the identified targets there were also factors involved in female sex development, such as on *FZD1* and *PBX1*. *Fzd1* encodes a WNT receptor whose expression is markedly induced in both mural granulosa cells and cumulus cells during the preovulatory period, in a manner similar to *Wnt4* (30). *WNT4* is a secreted protein essential for female sex development, in that it suppresses male sexual differentiation in mice and patients (31, 32). In humans, additional copies of *WNT4* have been associated with male-to-female sex reversal and *WNT4* loss of func-

tion causes DSD in women (33, 34). In a recent mouse study *Fzd1* has been showed to be required for normal female fertility, with *Fzd1* (-/-) females sterile or sub fertile (30). *PBX1*, which we found to be negatively regulated by CBX2, has been shown to play a role in female internal sex organ differentiation. In mice, the loss of *Pbx1* resulted in absence of Müllerian ducts (35). In the same context we observed that CBX2 negatively regulates *FOXL2* expression. *FOXL2* is considered to be the female counterpart of *SOX9* (36). In this regard, we hypothesize that CBX2 functions, at least in part, by actively repressing ovarian development in XY gonads which is in agreement with previous data showing that mutations in *CBX2* causes male-to-female reversal with ovarian tissue in *SRY*-positive mice (11) and in humans (5).

Interestingly, when we repeated the experiments by introducing a mutated form of *CBX2*, we previously described as being the molecular basis of 46, XY DSD human patient (5), there was no effect on all the selected targets compared to the control, confirming the loss-of-function nature of the diseases causing mutations and confirming that the observed effects on gene expression are *CBX2*-specific.

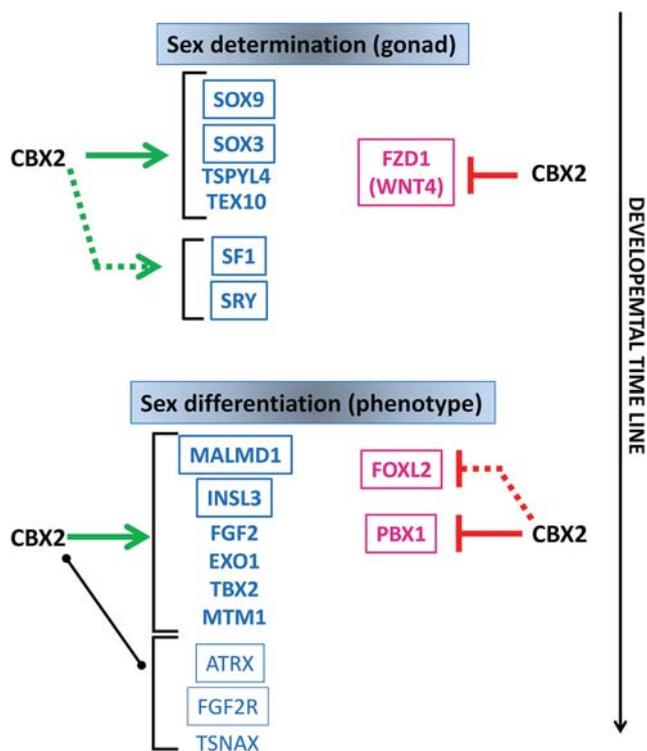
Taken together, our findings represent the first comprehensive genome-wide study of *CBX2* binding coupled with the expression profiling it further extends the role of *CBX2* in human sexual differentiation, and it also unravels the complexity and context-dependence of *CBX2*-modulated responses. In fact, it appears that *CBX2* on one hand acts in the testis by stimulating the expression of male specific genes, eg, *SOX9*, and simultaneously suppresses the female pathway by negatively regulating *FOXL2* and the *WNT4* pathway (Figure 6). The mechanisms underlying this bifunctional role are speculation, but are most likely due to differences in the composition of regulatory complexes (37) and confirm that, as nicely put by Sekido and Lovell-Badge “sex determination is a story of opposing forces and crucial alliances.[] It is a matter of timing (and expression level) that determines which pathway wins” (38).



**Figure 5.** Genes down-regulated in a *CBX2*-dependent manner. qRT-PCR quantification of mRNA levels for putative *CBX2* targets that are down-regulated by *CBX2*. NT2-D1 was treated as in Figure 4. *CBX2* overexpression (black) decreased the relative expression of *PBX1*, *FOXL2*, and *FZD1*, compared to cells transfected with empty vector (white). Upon *CBX2* knockdown (gray) the relative expression of *PBX1*, *FOXL2*, *FZD1*, were up-regulated upon comparing to scrambled siRNA (wide lines). Transfecting the cells with equimolar amounts of the mutated forms (P98L and R443P) of *CBX2* (narrow lines) showed no effect on the relative expression of all the target genes compared to control. The data in all graphs are the average of three independent experiments, error bars represent standard deviation from the mean and values are expressed as relative to control = 1. (\*\*\*\*)  $P < .0001$ ; (\*\*\*)  $P < .001$ ; (\*\*)  $P < .01$ ; (\*)  $P < .05$ .

Moreover, our study shed more light on novel and potential pathological mechanisms that could be involved in DSD: further characterization of how these new targets fit into an expanding CBX2-regulated network should reveal how CBX2 activation and suppression can impact our understanding of DSD pathogenesis and ultimately DSD diagnosis and management. Finally, since all these effects were abolished when the mutated CBX2 was tested, the hypothesized role of CBX2 as both a stimulator of male and inhibitor of female pathway might help to explain the presence of ovarian tissue in the gonads of our CBX2 deficient 46, XY DSD patient (5) and shed further light on the possible mechanism of the disease.

The future of clinical genetic analysis is doubtless next-



**Figure 6.** Potential role of CBX2 in the regulation of sex development. The developmental stages plotted in relation to time, are indicated as bipotential when both sexes are indistinguishable from one another (except for their karyotype), sex determination as the decision determined by several factors to develop into a male or a female gonad, and sex differentiation, the step determining the actual sexually dimorphic phenotype. In blue the male and in pink the female factors. The framed factors were found to be related to human disease. The framed factors were found to be related to human disease. Solid lines between CBX2 and its targets indicate a direct interaction (binding of CBX2 on target's DNA sequences). Green arrows indicate stimulatory effects of CBX2 on the indicated target; the red «t» indicates inhibitory effects. Broken lines indicate a functional but not physical interaction (changes in expression without direct binding). A black solid line indicates direct binding of CBX2 on the target without any effect on expression. Since all these effects were abolished when the mutated CBX2 was tested, the hypothesized role of CBX2 as both a stimulator of male and inhibitor of female pathway might help to explain the apparent presence of ovarian tissue in the gonads of our CBX2 46, XY patient (5).

generation sequencing (NGS). Nevertheless, whole exon sequencing (WES) identifies around 23%–25% of the genetic causes of diseases (39), and the percentage seems to be even lower in DSD (<10%, personal communication). Several reasons beside technical pitfalls might be accounted for these results. For example, the cause of the disease might be located outside the coding sequences, the disease is not monogenic and too many candidate variants remained after filtering or part of the mechanism of disease can be ascribed to differences in gene regulation rather than gene structure. Although WES remains a valuable tool in DSD, the study of regulatory networks is going to complement NGS in the understanding of these conditions. Experiments like ours leading to the discovery and understanding of pathways and networks will help “connecting the dots” in the interpretation of the NGS studies.

## Acknowledgments

We thank Dr. Bas van Steensel, Netherlands Cancer Institute for providing the DamID constructs.

Received October 28, 2014. Accepted December 26, 2014.

Address all correspondence and requests for reprints to: Prof. Anna Lauber-Biason, Division of Endocrinology, Department of Medicine, University of Fribourg, Chemin du Musée 5, 1700 Fribourg, Switzerland., Phone: +41 26 300 8534, E-mail: anna.lauber@unifr.ch.

This work was supported by Funding informations: the Swiss National Science Foundation Grant Nr. 320030 130645 to ABL.

Disclosure Summary: The authors have nothing to disclose

## References

- Ohnesorg T, Vilain E, Sinclair AH. The Genetics of Disorders of Sex Development in Humans. *Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation*. 2014;
- Baxter RM, Vilain E. Translational genetics for diagnosis of human disorders of sex development. *Annual review of genomics and human genetics*. 2013;14:371–392.
- Bashamboo A, Ledig S, Wieacker P, Achermann JC, McElreavey K. New technologies for the identification of novel genetic markers of disorders of sex development (DSD). *Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation*. 2010;4(4–5): 213–224.
- Hughes IA, Houk C, Ahmed SF, Lee PA, Lawson Wilkins Pediatric Endocrine Society/European Society for Paediatric Endocrinology Consensus G. Consensus statement on management of intersex disorders. *Journal of pediatric urology*. 2006;2(3):148–162.
- Biason-Lauber A, Konrad D, Meyer M, DeBeaufort C, Schoenle EJ. Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the CBX2 gene. *American journal of human genetics*. 2009;84(5):658–663.

6. Bracken AP, Helin K. Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nature reviews. Cancer*. 2009; 9(11):773–784.
7. Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. *Cell*. 2007;128(4):735–745.
8. Kennison JA. The Polycomb and trithorax group proteins of *Drosophila*: trans-regulators of homeotic gene function. *Annual review of genetics*. 1995;29:289–303.
9. Ringrose L, Paro R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual review of genetics*. 2004;38:413–443.
10. Sparmann A, van Lohuizen M. Polycomb silencers control cell fate, development and cancer. *Nature reviews. Cancer*. 2006;6(11): 846–856.
11. Katoh-Fukui Y, et al. Cbx2, a polycomb group gene, is required for Sry gene expression in mice. *Endocrinology*. 2012;153(2):913–924.
12. Biason-Laubert A. Control of sex development. *Best practice, research. Clinical endocrinology, metabolism*. 2010;24(2):163–186.
13. Katoh-Fukui Y, et al. Male-to-female sex reversal in M33 mutant mice. *Nature*. 1998;393(6686):688–692.
14. Katoh-Fukui Y, et al. Mouse Polycomb M33 is required for splenic vascular and adrenal gland formation through regulating Ad4BP/SF1 expression. *Blood*. 2005;106(5):1612–1620.
15. Eid W, et al. DNA end resection by CtIP and exonuclease 1 prevents genomic instability. *EMBO reports*. 2010;11(12):962–968.
16. Vogel MJ, et al. Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome research*. 2006; 16(12):1493–1504.
17. Orian A. Chromatin profiling, DamID and the emerging landscape of gene expression. *Current opinion in genetics, development*. 2006;16(2):157–164.
18. Butler JS, Koutelou E, Schibler AC, Dent SY. Histone-modifying enzymes: regulators of developmental decisions and drivers of human disease. *Epigenomics*. 2012;4(2):163–177.
19. Jakob S, Lovell-Badge R. Sex determination and the control of Sox9 expression in mammals. *The FEBS journal*. 2011;278(7):1002–1009.
20. Bhandari RK, Haque MM, Skinner MK. Global genome analysis of the downstream binding targets of testis determining factor SRY and SOX9. *PLoS one*. 2012;7(9):e43380.
21. Douglas NC, Heng K, Sauer MV, Papaioannou VE. Dynamic expression of Tbx2 subfamily genes in development of the mouse reproductive system. *Developmental dynamics : an official publication of the American Association of Anatomists*. 2012;241(2): 365–375.
22. Puffenberger EG, et al. Mapping of sudden infant death with dysgenesis of the testes syndrome (SIDDT) by a SNP genome scan and identification of TSPYL loss of function. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(32):11689–11694.
23. Wei K, et al. Inactivation of Exonuclease 1 in mice results in DNA mismatch repair defects, increased cancer susceptibility, and male and female sterility. *Genes, development*. 2003;17(5):603–614.
24. Small TW, et al. Wilms' tumor 1-associating protein regulates the proliferation of vascular smooth muscle cells. *Circulation research*. 2006;99(12):1338–1346.
25. Hu LJ, et al. Deletions in Xq28 in two boys with myotubular myopathy and abnormal genital development define a new contiguous gene syndrome in a 430 kb region. *Human molecular genetics*. 1996;5(1):139–143.
26. Kuroiwa A, et al. Additional copies of CBX2 in the genomes of males of mammals lacking SRY, the Amami spiny rat (*Tokudaia osimensis*) and the Tokunoshima spiny rat (*Tokudaia tokunoshimensis*). *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*. 2011;19(5):635–644.
27. Murata C, Yamada F, Kawauchi N, Matsuda Y, Kuroiwa A. The Y chromosome of the Okinawa spiny rat, *Tokudaia muenninki*, was rescued through fusion with an autosome. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*. 2012;20(1):111–125.
28. Papaioannou MD, et al. Loss of Dicer in Sertoli cells has a major impact on the testicular proteome of mice. *Molecular, cellular proteomics : MCP*. 2011;10(4):M900587MCP900200.
29. Cutting AD, et al. The potential role of microRNAs in regulating gonadal sex differentiation in the chicken embryo. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*. 2012;20(1): 201–213.
30. Lapointe E, et al. FZD1 regulates cumulus expansion genes and is required for normal female fertility in mice. *Biol Reprod*. 2012; 87(5):104.
31. Biason-Laubert A, Konrad D, Navratil F, Schoenle EJ. A WNT4 mutation associated with Mullerian-duct regression and virilization in a 46,XX woman. *The New England journal of medicine*. 2004; 351(8):792–798.
32. Vainio S, Heikkila M, Kispert A, Chin N, McMahon AP. Female development in mammals is regulated by Wnt-4 signalling. *Nature*. 1999;397(6718):405–409.
33. Jordan BK, et al. Up-regulation of WNT-4 signaling and dosage-sensitive sex reversal in humans. *American journal of human genetics*. 2001;68(5):1102–1109.
34. Biason-Laubert A. WNT4, RSPO1, and FOXL2 in Sex Development. *Seminars in reproductive medicine*. 2012;30(5):387–395.
35. Schnabel CA, Selleri L, Cleary ML. Pbx1 is essential for adrenal development and urogenital differentiation. *Genesis*. 2003;37(3): 123–130.
36. Veitia RA. FOXL2 versus SOX9: a lifelong “battle of the sexes”. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 2010;32(5):375–380.
37. Volkel P, Le Faou P, Vandamme J, Pira D, Angrand PO. A human Polycomb isoform lacking the Pc box does not participate to PRC1 complexes but forms protein assemblies and represses transcription. *Epigenetics : official journal of the DNA Methylation Society*. 2012;7(5):482–491.
38. Sekido R, Lovell-Badge R. Sex determination and SRY: down to a wink and a nudge? *Trends in genetics : TIG*. 2009;25(1):19–29.
39. Atwal PS, et al. Clinical whole-exome sequencing: are we there yet? *Genetics in medicine : official journal of the American College of Medical Genetics*. 2014;16(9):717–719.
40. Ogata T, Sano S, Nagata E, Kato F, Fukami M. MAMLD1 and 46,XY disorders of sex development. *Seminars in reproductive medicine*. 2012;30(5):410–416.
41. Moalem S, et al. XX male sex reversal with genital abnormalities associated with a de novo SOX3 gene duplication. *American journal of medical genetics. Part A*. 2012;158A(7):1759–1764.
42. Bagheri-Fam S, et al. Loss of Fgfr2 leads to partial XY sex reversal. *Developmental biology*. 2008;314(1):71–83.
43. Carmichael SL, et al. Hypospadias and genes related to genital tubercle and early urethral development. *The Journal of urology*. 2013;190(5):1884–1892.
44. Reardon W, Gibbons RJ, Winter RM, Baraitser M. Male pseudohermaphroditism in sibs with the alpha-thalassemia/mental retardation (ATR-X) syndrome. *American journal of medical genetics*. 1995;55(3):285–287.
45. Chennathukuzhi V, et al. Mice deficient for testis-brain RNA-binding protein exhibit a coordinate loss of TRAX, reduced fertility, altered gene expression in the brain, and behavioral changes. *Molecular and cellular biology*. 2003;23(18):6419–6434.