The developmental origins of sex-biased expression in cardiac development

Daniel F. Deegan¹, Reza Karbalaei²†, Jozef Madzo¹, Rob J. Kulathinal² and Nora Engel¹*†

Abstract

Background: Expression patterns between males and females vary in every adult tissue, even in organs with no conspicuous dimorphisms such as the heart. While studies of male and female differences have traditionally focused on the influence of sex hormones, these do not account for all the differences at the molecular and epigenetic levels. We previously reported that a substantial number of genes were differentially expressed in male and female mouse embryonic stem (ES) cells and revealed dose-dependent enhancer activity in response to Prdm14, a key pluripotency factor expressed more highly in female ES cells. In this work, we investigated the role of Prdm14 in establishing sex-specific gene expression networks. We surveyed the sex-specific landscape in early embryogenesis with special reference to cardiac development. We generated sex-specific co-expression networks from mouse ES cells, examined the presence of sex-specific chromatin domains, and analyzed previously published datasets from different developmental time points to characterize how sex-biased gene expression waxes and wanes to evaluate whether sex-biased networks are detectable throughout heart development.

Results: We performed ChIP-seq on male and female mouse ES cells to determine differences in chromatin status. Our study reveals sex-biased histone modifications, underscoring the potential for the sex chromosome complement to prime the genome differently in early development with consequences for later expression biases. Upon differentiation of ES cells to cardiac precursors, we found sex-biased expression of key transcription and epigenetic factors, some of which persisted from the undifferentiated state. Using network analyses, we also found that Prdm14 plays a prominent role in regulating a subset of dimorphic expression patterns. To determine whether sex-biased expression is present throughout cardiogenesis, we re-analyzed data from two published studies that sampled the transcriptomes of mouse hearts from 8.5 days post-coitum embryos to neonates and adults. We found sex-biased expression at every stage in heart development, and interestingly, identified a subset of genes that exhibit the same bias across multiple cardiogenic stages.

Conclusions: Overall, our results support the existence of sexually dimorphic gene expression profiles and regulatory networks at every stage of cardiac development, some of which may be established in early embryogenesis and epigenetically perpetuated.

Keywords: Sex chromosomes, Regulatory networks, Cardiogenesis, Sexual dimorphism, Differentiation

Background

It has long been acknowledged that clinical presentation of cardiovascular disease differs between men and women. Even in healthy adults, there are baseline sex differences in cardiovascular structure and function [1]. With the advent of sensitive sequencing technologies, a surprising amount of transcriptional and epigenomic variability has recently been shown between men and women in most adult tissues, including the heart [2–5]. Detailed studies in cardiomyocytes in humans, rats, and mice have also revealed sexual dimorphisms in transcriptome and function [6]. Most of these differences have been attributed to hormonal factors, yet results from many studies have shown that pathways other than hormones play an important role [7, 8]. For example, sex chromosomes contribute independently to sex biases in gene expression, although the

* Correspondence: noraengel@temple.edu
† Daniel F. Deegan and Reza Karbalaei contributed equally to this work.
² Fels Institute for Cancer Research, Lewis Katz School of Medicine, Temple University, 3400 N. Broad St, Philadelphia, PA 19140, USA
Full list of author information is available at the end of the article
specific sex chromosome-linked genes and their downstream targets have not been elucidated.

Genetic and epigenetic factors involved in normal cardiac development have been extensively characterized [9–14] and the transcriptional networks vital for cardiogenesis are well established. Generally, there has been no expectation in the developmental field that sex is relevant to early embryonic processes. Yet many congenital heart defects exhibit sex biases in presentation, mortality, and morbidity [15] and are primarily due to disruptions occurring before gonad formation. Furthermore, gestational insults, such as maternal undernutrition, are associated with sex-specific alterations in fetal heart development [16]. These imbalances have not been explained at either the genetic or developmental level and indicate that sex is an important biological variable during early embryogenesis.

In fact, sex-specific expression differences in early embryogenesis are widespread across the animal kingdom. Recent studies in non-mammalian model organisms have reported sex-biased expression at stages in which visible phenotypic differences between the sexes have not yet become apparent [17]. However, the question of whether this is also true of mammals has rarely been addressed.

One exception is the growing body of reports on mouse embryonic stem (ES) cells, which are self-renewing, pluripotent derivatives from pre-implantation embryos. Understanding the gene networks that control ES cells has been a major focus for many years [18–20] and recently, a surprising amount of sexual dimorphism in gene expression has been revealed in ES cells in both mice and humans [21–24]. Some expression differences were expected due to the presence of two active X chromosomes in females versus one in male cells. However, the majority of biases emanate from autosomal genes, including genes encoding dose-dependent transcription factors (TFs) and epigenetic and remodeling enzymes (EREs). This suggests that sex-specific gene networks are established by the sex chromosomes before X chromosome inactivation (XCI) occurs in female cells. Sophisticated network analyses have provided insight into the biology of organ development and can be deployed on the available data to address this possibility.

Evidence supports that sex-biased expression of regulatory factors in early embryogenesis establishes sex-specific epigenomic landscapes. Yet whether these differences are reversed by dosage compensation or are perpetuated during embryogenesis, with consequences for organogenesis and beyond, is unknown. Thus, it is necessary to characterize male and female transcriptomes over ontogeny in mammalian systems and to determine if they are connected to later adult phenotypes.

Here, we propose that the sex-biased expression of certain TFs and EREs in early development marks the genome with lasting effects across the lifespan [25]. We posit that while lineage specification decreases the range of sex-biased gene expression, sex-specific epigenetic marks persist and result in differential expression at later developmental stages [25]. We characterize candidates for these effects by elucidating co-expression and protein–protein interaction networks underlying the sex biases in male and female mouse ES cells. Our results highlight a co-expression module that is highly correlated with sex chromosome composition and identifies Prdm14, a sex-biased gene with higher expression in female ES cells, as a key regulator of sex biases in ES cells. Using heart development as a model process, we report sex-biased expression in male and female ES cells differentiated to cardiac precursors, in vivo embryonic hearts and in adult cardiomyocytes. By focusing on transcriptional and epigenetic factors, we identify a subset of sex differences established in early embryogenesis that persist throughout lineage determination and cardiac organogenesis. Furthermore, we find evidence that Prdm14 regulates target genes that are sex-biased during cardiac development and, surprisingly, in the adult heart, when Prdm14 is no longer expressed.

Methods

Construction of weighted gene co-expression network and modules

We used a previously published RNA-sequencing (RNA-seq) dataset from six male (40, XY) and six female (40, XX) mouse ES cell lines (GSE90516) for network analysis [24]. We derived these cell lines from independent F1 hybrid blastocysts resulting from reciprocal crosses of mouse substrains C57BL/6 and CAST/EiJ by natural mating. Each cell line was maintained in ES cell culture medium (DMEM, 15% fetal calf serum, 1 mM sodium pyruvate, 2 mM l-glutamine, 1% non-essential aminoacids, 0.1 mM 2-mercaptoethanol and 1000 U/ml leukemia inhibitory factor) in 5% CO2 at 37°. The data was generated using HiSeq 2500 single end reads of 50 base pairs. We reported hundreds of coding and non-coding RNAs that were differentially expressed between male and female ES cell lines, after filtering for strain-specific effects [24].

To prevent bias, all aligned transcripts were used to establish a weighted gene co-expression network analysis (WGCNA), a widely used systems biology method that employs gene expression data to construct a scale-free network [26]. The WGCNA package in R, version 1.6, is available at https://cran.r-project.org/web/packages/WGCNA/index.html. For the analysis herein, Pearson’s correlation matrices were calculated for all pairs of genes evaluating the correlation coefficient between gene m and gene n such that Smn = |cor(m,n)|. Next, the Pearson’s correlation matrices were transformed into matrices defining
connection strengths using the power function \( a_{mn} = \text{power}(S_{mn}) = |S_{mn}|^\beta \). In doing so, strong correlations are emphasized and the influence of weak correlation is reduced on an exponential scale. To obtain a scale-free network, we performed network topology analysis for thresholding powers from 1 to 20. The lowest power value for scale-free topology was 10, so \( \beta \) was set at 10.

The connectivity of pairs of genes was evaluated by calculating topology overlap (TO). TO is a robust indicator of the relationships among neighborhoods of genes. The TO was then used to perform hierarchical average linkage clustering to identify gene co-expression modules. Modules are branches of a hierarchical cluster tree defined using the top-down dynamic tree cut method [27] with a minimum module size of 50 genes. After module identification, a \( t \) test was used to calculate the \( p \) value of candidate genes. The gene significance (GS) was defined as the mediator value of each gene (\( GS = \log P \)). From this, the module significance (MS) was defined using the average GS from all the genes within said module.

**Transcription factor motif analysis**

The set of genes within the module most highly correlated with cell sex, comprised of 1624 genes, was analyzed for known and de novo transcription factor motif binding sites. The parameters were set to cover the promoter using \(-5000\) to \(+1000\) bp of the transcriptional start site in the HOMER online software suite (http://homer.ucsd.edu/homer/) [28].

**Ingenuity pathway analysis**

We analyzed gene sets using the ingenuity systems pathways analysis (IPA) tool (Qiagen, Redwood City, CA). Datasets were subjected to IPA Core Analysis and then analyzed using IPA Upstream Regulator, Downstream Effects, and Canonical Pathways analytic tools. To capture regulatory networks, we focused on transcription factors and epigenetic and remodeling enzymes. The IPA output was exported as Microsoft Excel files to prepare the supplemental tables.

**Chromatin immunoprecipitation and sequencing**

Four low-passage (p7–9) independent mouse ES cell lines, two male lines (40, XY) and two female lines (40, XX), were grown on inactivated C57BL/6 mouse embryonic fibroblasts (MEFs). The MEFs are prepared from pooled embryos and include both male and female cells. ES cells were passaged at least twice prior to harvest to achieve high cell numbers for chromatin immunoprecipitation and sequencing (ChIP-seq). Cells were collected using 0.25% Trypsin + EDTA and MEF-depleted for 1 h at 37 °C in 5% carbon dioxide. After collecting the ES cells, residual MEFs were less than 1.5% of the final cell suspension. Because of their low numbers and the fact that they are a mixed population of male and female cells, any remaining MEFs are not expected to skew the results obtained from the ES cells. ES cells were cross-linked using formaldehyde at a final concentration of 1% followed by quenching with 1 M glycine. Sonication, immunoprecipitation, library construction, and sequencing were performed as previously described with minor modifications [29]. Briefly, three consecutive lysis buffers were used to ensure adequate nuclei release. Sonication was performed in a Q-Sonica cup horn sonication system using amplitude 70 with 30 s ON/OFF cycles for 10–15 min dependent on desired size and sonication efficiency of each sample. Samples were sonicated to a range of 100–500 base pairs. Sonicated chromatin was diluted in immunoprecipitation buffer. Two million cells were used for each immunoprecipitation (IP), with five consecutive IPs for each histone modification of interest. Ten percent of the initial sample volume per IP was set aside to serve as input control prior to addition of the appropriate antibody. Additional file 1: Table S1 provides the specifics for the antibodies used, with 2.5 μg of each antibody per IP. To isolate the antibody-bound fragments of interest, we used 50 μl of a 50/50 mix of Dynabeads TM Protein A (catalog # 10002D, lot # 00448844) and Protein G (catalog # 10003D, lot # 00486042) and an overnight incubation at 4 °C.

Beads were washed with RIPA buffer for five consecutive washes followed by a single wash with Tris-EDTA buffer. Complexes were eluted from the beads with 50 mM Tris, 10 mM EDTA, 1% (w/v) SDS, pH 8.0, and crosslinks were reversed. Concentration of resultant DNA was determined using Qubit according to manufacturer's protocol. ChIP-seq libraries were prepared using DNA SMARTTM ChIP-Seq Kit. Sequencing was performed using Illumina HiSeq 2500 generating single end reads of 50 base pairs. Sequences were aligned to mouse genome assembly (mm9) using Bowtie2 v2.1.0 with default settings [30]. Using the Bedtools software suite for genome arithmetic [31], we determined the degree of enrichment of reads genome-wide. For data visualization, we used R software suite. To visualize enrichment patterns of histone modifications at promoters and enhancers, we used ngs.plot [32], using enhancer annotations from a previous report [33].

**Protein-protein interaction networks**

Protein-protein interaction networks (PPIs) were constructed with the STRING database using all differentially expressed genes between male and female ES cells (STRING version 10.5 [34]). To improve the quality of the resulting network, the “minimum required score” option was set to 0.7 and “text mining resources” were ignored. Analysis and graphing of the networks were performed in the Gephi software (version 0.9.2) [35].
Functional modules were detected using an algorithm to partition the network into communities of densely connected nodes [36]. Gene ontology (GO) analysis was performed using the ClueGO plug-in from Cytoscape [37]. GO terms were summarized using the REVIGO website (http://revigo.irg.hr/revigo.jsp) [38]. Network topology analysis and selection of important genes were done as previously described [39, 40].

**ES cell differentiation**

Two of the male and female ES cell lines from which ChIP-seq was performed were subjected to a standardized differentiation protocol that directs the stepwise differentiation of early embryonic cells into cardiac precursors [41], as assayed by marker gene analysis. Cells were cultured with leukemia inhibitory factor (LIF) on mouse embryonic fibroblasts (MEFs). Before differentiation, ES cells were dissociated, MEFs were eliminated as detailed above, and embryoid bodies were derived by hanging drop culture in medium without LIF. After 4 days, embryoid bodies were harvested and grown in medium containing Activin A, BMP4, and VEGF as monolayers until beating foci were observed. This optimized protocol yields >75% cardiomyocytes [41]. At day 13 of initial LIF withdrawal, we picked beating foci from the plates and obtained RNA.

qPCR was performed to determine expression of pluripotency markers *Nanog* and *Oct4* and cardiomyocyte markers *Myh6* and *Tnnt2* on cDNA generated using SuperScript™ II (Invitrogen) and relative expression was assessed using PowerUp SYBR Green Master Mix (Thermo Fisher) and normalized to *β*-actin on the Applied Biosystems StepOnePlus Real-Time PCR System. RNA-seq was performed as previously described [24].

**Meta-analysis of publicly available data**

We leveraged existing expression datasets from across cardiac development in the mouse and stratified the data by sex when necessary. These collated data allowed us to determine whether there is dynamic sex-biased expression across cardiogenesis. Additional file 2: Table S2 details all the datasets examined herein.

Single cell data for 8.5, 9.5, and 10.5 days post coitum (dpc) embryonic [42] and neonatal mouse hearts [43] were downloaded and processed as follows: (1) if the fragments per kilobase of exon per million reads mapped (FPKM) was <1, the gene was designated as not expressed; (2) genes with zero variance across all cells were removed. Cells were then sexed by determining the ratio of *Xist* to *Eif2s3y*, two oppositely biased genes, on a cell-by-cell basis. Cells with *Xist/Eif2s3y* ratios of at least 1.5 were considered as female and ratios below 1 were taken as male. t test analysis was performed on the samples from each stage, p values were used to calculate the false discovery rate (FDR), and genes with adjusted p value < 0.05 were selected as differentially expressed genes. Data for adult mouse hearts was already stratified by sex [44].

**Transcription factor binding analysis**

To detect recognition motifs of candidate transcription factors (TFs) in genes enriched in male or female cells, we used the genome-wide position matrix scanner from the Computational Cancer Genomics website (https://ccg.epfl.ch/pwmtools/pwmscan.php) with the JASPAR core vertebrate motif library (version 2018). We searched for *Lef1* MA0768.1 and *Zeb1* MA0103.3 motifs with a p value cutoff of 0.00001 with the Contra v3 tool (http://bioit2.irc.ugent.be/contra/v3/#/step/1) and uploaded the results as custom tracks in the UCSC browser.

**Results**

**Defining a gene network associated with sex-biased gene expression**

There is a substantial number of differentially expressed genes in male (40, XY) and female (40, XX) mouse embryonic stem (ES) cells, including transcription factors (TFs), and epigenetic and remodeling enzymes (EREs) [22–24]. Yet both male and female ES cells are pluripotent and can contribute to normal development. Thus, while overall pluripotency networks govern both XX and XY ES cells, we hypothesized that differentially expressed genes can shift network architecture or constitute subnetworks with distinct gene-gene correlations.

To determine whether the genes differentially expressed in XX and XY ES cells constitute sex-specific co-expression networks and identify genes with higher connectivity in each sex, we used normalized RNA-sequencing (RNA-seq) data from six male and six female mouse ES cell lines to perform weighted gene co-expression network analysis (WGCNA) [24, 27, 45] (see “Methods” section). Weighted gene co-expression network analysis allows partitioning of genes into modules that correlate with biological function and identifies the genes that are most likely to be crucial in regulating that function. WGCNA has been successfully applied to dissect the role of hormonal and sex chromosome effects in sex-biased co-expression networks in adult tissues [46].

Figure 1 shows the clustering dendrogram of co-expressed genes resulting from the WGCNA with the lowest power value for scale-free topology, β, set at 10. To avoid bias by pre-selecting genes with differential expression levels in male and female ES cells, we based the clustering on all aligned transcripts. Genes with similar patterns of expression were grouped into modules by hierarchical average linkage clustering using topological overlap [26]. The initial dynamic tree cut was further merged, to generate a subset of 11 distinct co-expression modules.
The first principal component of a given module is the module eigengene (ME), which represents the gene expression profile within that particular module. To understand the functional significance of the modules, we correlated the 11 MEs generated within the clustering dendrogram with traits of interest and isolated the most significant associations (Fig. 2a). According to the heatmap of module-trait correlations, sex showed a strong and independent association with a particular eigengene, the ME blue/violet ($r = 0.85$, $p = 5e^{-04}$) and consisted of 1624 genes, including 84 TFs and 43 EREs (Additional file 3: Dataset S1).

To validate the gene cluster with a separate method, we produced a hierarchical clustering heatmap using the expression levels of the 1624 genes in the blue/violet module. Input of the genes contained within the blue/violet module into this separate pipeline did indeed show separation of the mouse ES cell lines by sex (Fig. 2b).

**Distinct upstream regulators are associated with sex-biased functional pathways**

To identify regulatory pathways for the genes in the blue/violet module (Fig. 2), i.e., the module best correlated with sex, we performed ingenuity pathway analysis independently on the XX- and XY-enriched TFs and EREs (Additional file 4: Dataset S2). We found that the top pathway for XX-enriched TFs and EREs was “DNA methylation and transcriptional repression” ($p = 7.81 \times 10^{-4}$), with Max and Mycn as the top upstream regulatory molecules. Analysis of the XY-enriched TFs and EREs from the blue/violet module identified “Jak1 in Interferon Signaling” as the top pathway ($p = 2 \times 10^{-3}$). Top upstream regulators were predicted to be Irf9 and Npc1.

**Prdm14 motifs are enriched in promoters of sex-biased genes**

We asked if the sex-biased clustering of the blue/violet genes was being driven by specific transcription factors and reflected sex-specific regulatory networks. To test this, we used HOMER to identify known transcription factor binding sites within the gene set in the blue/violet ME [28]. HOMER motif analysis yielded significantly enriched TF motifs in the promoters of genes in the blue/violet module eigengene (Table 1). The transcription factor TEAD (TEA/ATTS domain) was the top and most significantly enriched motif ($p$ value $1e^{-15}$). TEAD proteins are pivotal transcription factors implicated in development as well as in cancer [47]. Leukemia inhibitory factor, present in the culture medium, activates the Yes-
associated protein (YAP) and TEA domain TEAD2 transcription factor pathway, which contributes to mouse ES cell maintenance of pluripotency and self-renewal. The stem cell factors Nanog and Oct3/4 are targets of the TEAD-pathway [48]. These pluripotency factors had similar expression levels among all male and female ES cell lines tested, and TEAD was not differentially expressed at the RNA level. However, it was previously reported that Tead1 and Tead2 are male-biased at the protein level [23]. Thus, further investigation is required to determine if these factors contribute to the sex-specific effects or whether they appear with the HOMER analysis due to their contribution to pluripotency per se.

Interestingly, Prdm14 is a top hit and the second highest hit in HOMER (Table 1). Prdm14 is expressed more highly in female (XX) than in male (XY) ES cells, a bias that occurs independently of whether the ES cells are cultured in LIF/serum or 2i [22, 24] and is also seen at the protein level [23]. Prdm14 is a bi-functional TF with a cardinal role in ES cell pluripotency and in the establishment

![Fig. 2 Relationships of consensus modules (module eigengenes) with sex.](image)

**Fig. 2** Relationships of consensus modules (module eigengenes) with sex. a Each row in the table corresponds to a consensus module identified by distinct colors along the left y-axis. Each module eigengene (ME) was evaluated in relationship to sex. Numbers in the table report the correlation of the corresponding ME with sex, with the p values shown in parentheses. The degree of correlation, positive and negative is provided by the colored scale on the right y-axis. b Clustering of the mouse ES cell lines based upon the module eigengene, blue/violet. Heatmap showing separation of the lines by sex chromosome complement (XY, male; XX, female; XO, X chromosome monosomic) when the 1624 genes, contained within the blue/violet module from WGCNA were evaluated.

**Table 1** HOMER motif analysis of the promoters of genes in the blue/violet module eigengene

| TF    | Consensus Sequence | p-value | % of targets |
|-------|-------------------|---------|--------------|
| TEAD  | TGCAAGCTC         | 1e-15   | 30.21%       |
| PRDM14| GGCGCTAGAGA       | 1e-14   | 20.47%       |
| Otx2  | TTTAATCTTACCA     | 1e-14   | 8.75%        |
| Egr1  | TACACCCCAACA      | 1e-12   | 0.86%        |

Genomic regions for −5000 bp to +1000 bp around the transcriptional start site were pulled for each of the 1624 genes contained within the blue/violet module eigengene and input into HOMER. Shown here are the four top hits and corresponding p values and percentage of coverage within target sequences.
of primordial germ cells. Prdm14 can either activate or repress gene expression, depending on its interacting partners [49]. Recruitment of polycomb repressive complex 2 (PRC2) by Prdm14 results in transcriptional repression, whereas cooperation with estrogen-related receptor β (Esrβ) activates target gene expression. However, the mechanisms by which Prdm14 selectively partners with its alternative co-factors, resulting in gene activation or repression, are not understood. Nevertheless, Prdm14 is a strong candidate for regulating gene expression differentially in male and female ES cells and establishing sex-biased epigenetic marks.

Prdm14 target genes encoding TFs have sex-biased expression
To identify downstream targets of Prdm14, we curated and compared publicly available expression profiles of ES cells depleted of Prdm14, focusing on TFs and EREs. Several studies have reported Prdm14 knockout or knockdown in ES cells, with inconsistent results, likely due to varying culture conditions, strains, and karyotypes [50–52]. Therefore, we focused on a report with siRNA-mediated knockdown of Prdm14 in wild-type female 129/Ola ES cells, with the caveat that culture conditions were 2i–depleted female ES cells, such as Mitf, Zeb1, and Prdm14 itself, are enriched in wild-type female ES cells. More than 10% of genes followed this pattern. This confirms that Prdm14 regulates a subset of genes, while also indicating that there are other factors involved in sex-biased expression.

Male and female ES cells exhibit sex biases in chromatin modifications
To determine if the differential transcriptomes between XX and XY ES cells are reflected in the chromatin structure, we performed chromatin immunoprecipitation and sequencing (ChIP-seq) on six early passage independent ES cell lines of each sex, i.e., the same cell lines for which we had reported sex-biased expression [24]. Antibodies against histone modifications H3K4Me1, H3K27Me3, and H3K27Ac were used to precipitate chromatin substrates with our standard protocol. The presence of H3K27Ac, indicating active chromatin, showed a significant difference between XX and XY ES cells at known enhancer regions (Fig. 3). This suggests that the main biases between XX and XY ES cells are established by TFs and EREs that bind and modify enhancer sequences.

To determine whether there was concordance between Prdm14 binding, gene expression biases, and differential chromatin modifications, we integrated available Prdm14 ChIP-seq data in ES cells [50] with our sex-specific chromatin studies for genes that respond to Prdm14 according to the knockdown studies.

Our analysis identified three groups of differentially expressed genes: (1) genes that exhibited sex-biased chromatin modifications and Prdm14 binding, (2) genes with Prdm14 occupancy and no sex-specific histone modifications, and (3) genes with neither detectable sex-biased chromatin modification nor Prdm14 occupancy. For example, Dmnt3l, more highly expressed in male ES cells, has a Prdm14 binding site 40 kb downstream of the transcript start site, which is enriched in H3K27Me3, a repressive mark, in female ES cells (Fig. 4). One of the Prdm14 binding sites downstream of Mitf, more highly expressed in XX ES cells, has enrichment of H3K27Ac in those cells. Hoxb9 shows a similar pattern, with a Prdm14 binding site enriched in H3K27Ac in female ES cells which have higher expression. On the other hand, there are several Prdm14 binding sites upstream and in the promoter of Meis2, but we did not detect differential histone modifications in male and female ES cells, although it is more highly expressed in female cells. Genes such as Sohlh2 do not have apparent Prdm14 binding in their vicinity, indicating that they are regulated by other, as yet unknown TFs.

Protein-protein interaction network analysis of ES cell transcriptomes reveals overlap with Prdm14 target genes
Data from the differentially expressed genes in male and female ES cells were used to construct a protein-protein interaction (PPI) network (Fig. 5a). Overlaying the information from the sex differences in gene expression shows that there are sex-biased modules within the global interaction network. We compared the genes from the blue/violet module eigengene from the WGCNA to the nodes of the PPI network. Two hundred twenty-five genes were shared between them (green nodes in Fig. 5b). Analysis of the network revealed six modules (Additional file 6: Figure S1), one of which contained the most important nodes based on topological analysis (degree, betweenness, and closeness centrality metrics in Additional file 5: Dataset 3). GO analysis of this module showed “blood vessel morphogenesis” and “Bmp signaling” as top terms (Additional file 7: Dataset S4).

Prdm14 was contained in the module with the most important nodes and showed connections to Dazl, Tcl1, Wnt3, Cdx2, Dmnt3b, Prdm6, Bmp4, Lin28a, Lefty2, T, and Gata4. Strikingly, many of the nodes in this PPI module, such as Zeb1, Lefty1, Gata4, Dusp6, and Sox11, are direct Prdm14 transcriptional targets in ES cells.
Male and female cardiac precursors also show sex biases in gene expression

Thus far, we showed that there are sex-specific expression and protein-protein interaction networks in ES cells. Upon differentiation of female ES cells, one of the two X chromosomes is inactivated, a massive epigenetic event that equalizes most of the X-linked genes between males and females. This transition mirrors the in vivo process of blastocyst implantation, during which female embryos undergo X chromosome inactivation (XCI).

To determine whether some sex-biased expression differences were perpetuated after XCI and during the beginning stages of lineage determination, we subjected two male and female ES cell lines to an optimized differentiation protocol to generate cardiac precursors and performed RNA-seq (Fig. 6). At day 13 after LIF withdrawal, ES cells have differentiated to cardiac precursors corresponding to 8.5–9.5 days post coitum (dpc) cardiac progenitors in vivo. RT-PCR confirmed that stem cell markers such as Nanog and Oct4 were downregulated, whereas markers of cardiac differentiation, such as Tnnt2 and Myh6, were upregulated in both sexes, as previously reported (Additional file 8: Figure S2) [10, 53].

We compared transcriptomes between differentiated male and female cell lines and found 157 genes that were differentially expressed at a FDR < 0.01 (Additional file 9: Dataset S5). The Xist non-coding RNA, which is involved in X chromosome inactivation, was more highly expressed in female cells, as expected. The male cells showed higher expression of 2 Y chromosome-linked genes, Ddx3y and Uty (Kdm6c). Interestingly, four TFs were more highly expressed in male cells, Ferd3l, Pou3f3, Six6, and St18. Ferd3l and Pou3f3 have nearby Prdm14 binding sites in undifferentiated ES cells, although we did not detect differential histone modifications in their vicinity (Additional file 10: Figure S3). Overall, this data shows that although the number of genes exhibiting sex differences diminishes during lineage determination, some biases persist. ChIP-seq data for cardiac precursors derived from male and female ES cells are needed to determine which epigenetic differences also persist after differentiation.

Sex biases in cardiac expression exist in early cardiac developmental stages in vivo

To elucidate how sex biases in gene expression vary during cardiac development, we collated and analyzed single cell transcriptional profiles from mouse embryonic hearts at 8.5, 9.5, and 10.5 dpc [42] and post-natal day 1 (p1) (Additional file 2: Table S2) [43]. Single cell data was downloaded and sexed (Additional file 11: Dataset S6). We found that there were hundreds of sex-biased genes at every stage. Some of these were stage-specific and some were common to two or more time points. For example, Lef1 was more highly expressed in male than female ES cells, and the same was true for 8.5 dpc and p1 hearts. Tbx20 was also enriched in male ES cells, cardiac precursors, and in 10.5 dpc and p1 hearts.

The majority of genes with sex-biased expression were male-biased at every stage. The number of female-enriched genes peaked dramatically at 9.5 dpc and decreased thereafter. At 8.5 dpc, only three X-linked genes, including Xist, were female-biased, whereas 19 X-linked genes were male-biased. At 10.5, eight X-linked genes were more highly expressed in females, including Xist, Tsix, and three genes that had not been characterized as escapees. More than 30 X-linked genes showed male-biased expression, indicating that some genes are not dosage compensated by X chromosome inactivation, at least at this stage in this tissue.

Protein-protein interaction networks were constructed with the sex-stratified expression data from 8.5, 9.5, and
10.5 dpc hearts (Fig. 7, Additional file 12: Figure S4, Additional file 5: Dataset S3). Sex biases in specific modules varied across the developmental stages, suggesting a highly dynamic but constant pattern of sexual dimorphism at the molecular level.

**Adult male and female hearts have sex-specific pathways**

To explore whether there are expression differences in male and female adult C57BL/6 mouse hearts, we inspected recently published transcriptomic data across 17 tissues, stratified by sex [44]. Strikingly, 908 and 148 genes exhibited expression biases in adult male and female hearts, respectively, again showing that male-biased genes are more numerous. Interestingly, 38 X-linked genes were male-biased, suggesting male-specific regulation of these genes.

We investigated whether TFs that were sex-biased in adult hearts showed Prdm14 binding in ES cells. We found that *Nkx2.5, Lef1, Id2, Ikzf3*, and *Srebf2* had Prdm14 occupancy in or near their promoter regions (Additional file 13: Figure S5), suggesting that their sex differences could have been established in early development. Differential histone modifications, however, were not apparent in ES cells in these regions.
We used Ingenuity Pathway Analysis to identify enrichment of biological network components in the sex-specific gene signatures in the adult heart. The top canonical pathways differed between male and female cardiac cells (Additional file 14: Datasets S7 and S8). Cardiovascular disease was the top disease association and cardiovascular system development and function was one of the top networks in importance for females, but surprisingly, cancer was the top disease association as well as the top network for males. Regulatory component analysis predicted distinct upstream regulatory factors for the male and female expression patterns. For example, Tp53, Nr3c2, and Tbx5 were among the top transcriptional regulators for female cells, whereas Ncor1 and Smad3 were identified for male cells.

**Conserved sex-biased expression between mouse and human hearts**

We compared sex-biased genes in adult heart ventricles between mouse and human. Differentially expressed genes between male and female human hearts were obtained from DeMeo et al., in which expression from the GTEx portal was stratified by sex [4]. There are 70 and 328 genes that are enriched in females and males, respectively, in both mouse and human (Additional file 15: Dataset 9). Among these are TFs Bhlhe40, Tcf15, Npas3, and Mafa, which are enriched in female hearts. Males show higher levels of Ehf, Etv1, Foxk1, Ikzf2, Meis2, and Tbx20, among others and of EREs Hat1, Cdy1, and Rad51l2.

**There is sex-biased expression at specific developmental time points for important cardiac regulators**

To query temporal changes in sex-biased expression profiles, we compared differential expression from ES cells (our data), embryonic and neonatal hearts, and adult cardiac myocytes [44]. Figure 8 and Tables 2 and 3 show sex-biased expression of TFs and EREs at each stage of heart development. Several different patterns can be visualized. Some genes encoding TFs and EREs are only expressed at one stage and others across several stages. For the latter group, there are subsets of genes that either maintain, acquire, lose, or even reverse their bias. A distinct group of genes, for example, Carhsp1 (male-biased) and Bhlhe40 (female-biased), display sex differences before gonad formation and the appearance of sex hormones. Our data also reveals sex disparities in expression that only become apparent in neonates and adults, suggesting that these respond, at least in part, to hormonal differences.

Thirty-six genes have conserved sex-biased expression in ES cells and adult hearts. Of those, six genes are more highly expressed in females at both stages, four of which are X-linked. Interestingly, only one of the X-linked genes has been previously described as escaping X chromosome inactivation (XCI) (Kdm6a) [54]. Thirty genes are male-biased in both ES cells and adult cardiomyocytes, including the three transcription factors Nfkb2, Lef1, Id2, and the epigenetic enzymes Uty and Prdm6.

Some genes expressed at early stages are still expressed in neonates or adults but lose their sex differences or even exhibit reversals in sex biases. In females, the X-linked Aff2 and Atrx lose their bias, which likely reflects
dosage compensation after X chromosome inactivation in female cells. However, Meis2 and Zfp9 switch to male-biased expression in adults (Table 2). Seven male-biased genes, including Irf8, Pbx2, Gata4, and Hdac5, which exhibit higher expression in male ES cells, become equally expressed in adult hearts of both sexes. Dot1l and Zfp296 reverse their bias and are more highly expressed in females at later stages (Table 3).

We also find several genes that are not expressed differentially in male and female ES cells and later acquire a sex bias. These are good candidates for genes regulated by hormonal factors, although Esr1, the only estrogen receptor expressed in the heart, is not differentially expressed between males and females. RNA encoding the androgen receptor is also not sex-biased in the adult heart, suggesting that hormonal regulation depends on other co-factors and/or differential chromatin environment of the target genes.

To further explore the role TFs expressed in early development have at later stages, we identified binding sites for Lef1 and Zeb1 within the regulatory regions of genes differentially expressed between male and female cardiomyocytes. Lef1 is enriched in male ES cells, 8.5 dpc embryonic hearts, and in neonatal and adult hearts. Genes that harbored Lef1 binding motifs included other TFs that are also male-biased in ES cells, such as Mixl1, Mesp1, Irf8, and Tbx20, but also genes that are only later expressed in the adult heart, such as Gata5 and Foxo6, which are also male-enriched (Fig. 9a). Zeb1 is enriched in female ES cells and not detected at later stages, but its cognate motifs are present in genes that are female-biased in adult heart, such as Cecr2 and Nkx2-5 (Fig. 9b). These results suggest that TFs expressed in early development may determine sex-biased gene expression at later stages.

Discussion

This study challenges the expectation that sex biases in gene regulation are non-existent during early development.
mammalian development. Although sex determination has been traditionally associated with processes leading to distinct reproductive systems in males and females, we show that sex biases appear soon after fertilization and may have sex-specific repercussions during organogenesis, some of which persist in adults. In the absence of time-series experiments from pre-implantation embryos throughout lineage determination and organogenesis, we capitalized on our own data and a series of previously published RNA-seq datasets.

Gene co-expression network analysis identifies Prdm14 as a key determinant of sex-biased gene expression in ES cells

Previous reports have identified thousands of genes differentially expressed in male and female ES cells and pre-implantation embryos in rodents, bovine, primates, and humans [21–24, 54–59]. In this work, we asked whether sexual dimorphism is detectable at the molecular level of gene expression and enriched in protein-protein interaction networks in early development. Both WGCNA and PPI networks revealed that important

Fig. 7 Protein-protein interaction networks in early cardiac development. PPIs were constructed from differentially expressed genes in 8.5, 9.5, and 10.5 dpc hearts as assayed by single-cell RNA-seq. Networks include sex-biased modules highlighted by red (female-enriched) and blue (male-enriched) nodes (based on data from Li, G. et al.)

Fig. 8 Expression of sex-biased transcription and epigenetic factors throughout development. Schematic heatmap representation of Tables 2 (a) and 3 (b) indicating expression and sex biases of transcription and epigenetic factors at each time point. Data was compiled from a female and b male ES cells, derived cardiac precursors (CP), hearts from 8.5, 9.5, and 10.5 days post coitum (dpc) embryos, neonates (p1), and adult mice (Ad). Each row is a specific transcription or epigenetic factor, with a total of 60 for females and 61 for males; the color denotes expression detected and enrichment in XX (red), XY (blue), or not biased (yellow). Group I: biased in ES cells, not expressed thereafter; groups II, III: biased in ES cells and same (II) or different (III) bias at other stages; group IV: biased after implantation but before gonadogenesis; group V: biased only after gonadogenesis
| Transcription factors and epigenetic remodelling enzymes | ES | CP 8.5 | 9.5 | 10.5 | p1 | Adult Chr | Name |
|---------------------------------------------------------|----|--------|-----|------|----|-----------|------|
| Hmgb3                                                   | XX |        |     |      |    |           | X    |
| Dmrtb1                                                  | XX |        |     |      |    |           | 4    |
| Hoxb9                                                   | XX |        |     |      |    |           | 11   |
| Hoxc8                                                   | XX |        |     |      |    |           | 15   |
| Klf8                                                    | XX |        |     |      |    |           | X    |
| Mitf                                                    | XX |        |     |      |    |           | 6    |
| Prdm14                                                  | XX |        |     |      |    |           | 1    |
| Rhox1                                                   | XX |        |     |      |    |           | X    |
| Sohlh2                                                  | XX |        |     |      |    |           | 3    |
| Spic                                                    | XX |        |     |      |    |           | 13   |
| Tbxi5                                                   | XX |        |     |      |    |           | 3    |
| Zeb1                                                    | XX |        |     |      |    |           | 18   |
| Zfp182                                                  | XX |        |     |      |    |           | X    |
| Zfp275                                                  | XX |        |     |      |    |           | X    |
| Zfp449                                                  | XX |        |     |      |    |           | X    |
| Zfp59                                                   | XX |        |     |      |    |           | 7    |
| Zxdh                                                   | XX |        |     |      |    |           | X    |
| MeCP2                                                   | XX |        |     |      |    |           | X    |
| Nkap                                                    | XX |        |     |      |    |           | X    |
| Ogt                                                     | XX |        |     |      |    |           | X    |
| Top2b                                                   | XX |        |     |      |    |           | 14   |
| Trim16                                                  | XX |        |     |      |    |           | 11   |
| ZmyM3                                                   | XX |        |     |      |    |           | X    |
| Apobec2                                                 | XX |        |     |      |    |           | 17   |
| Bax2b                                                   | XX |        |     |      |    |           | 2    |
| Prrx1                                                   | XX | XX     |     |      |    |           | 1    |
| Kdm6a                                                   | XX | XX     | XX  |      |    |           | X    |
| Lbh                                                     | XX |        | XX  |      |    |           | 17   |
| Meis2                                                   | XX |        | XY  |      |    |           | 2    |
| Zfp9                                                    | XX |        | XY  |      |    |           | 6    |
| Aff2                                                    | XX | XX     | nb  |      |    |           | X    |
| Atrx                                                    | XX |        | XY  | nb   |    |           | X    |
| Arid1b                                                  | nb | XX     | XX  | nb   |    |           | 17   |
| Bhlhe40                                                 | nb | XX     | XX  |      |    |           | 6    |
| Zfp51                                                   | nb | XX     | XX  |      |    |           | 17   |
| KiF15                                                   | nb | XX     |     |      |    |           | 6    |
| Mafk                                                    | nb | XX     | XX  |      |    |           | 5    |
| Aff3                                                    | nb | XX     | XX  | XY   |    |           |      |
| Heyl                                                    | XX |        |     |      |    |           | 4    |
| Hopx                                                    | XX |        |     |      |    |           | 5    |
| Dlx1                                                    | XX |        |     |      |    |           | 2    |
modules associated with sex are enriched in genes with Prdm14 cognate binding sites and are Prdm14 target genes. Prdm14 is important for pluripotency in ES cells [49, 52] and is a key regulator of primordial germ cell specification [60, 61]. In contrast to other PRDM family members, Prdm14 does not exhibit histone methyltransferase, but has been shown to partner with enzymes that catalyze post-translational modification of histones [49]. In fact, as seen in our ChIP-seq data, male and female ES cells have differential chromatin modifications, some of which are associated to Prdm14 occupancy at regulatory sequences. In addition, Prdm14 binding is found in promoters or neighboring regions of genes that are not expressed in ES cells. Thus, epigenetic marks established in pre-implantation stages can potentially result in sex-biased gene expression later in development.

Prdm14 expression is downregulated after differentiation of male and female ES cells and after implantation in vivo. However, female ES cells are developmentally delayed relative to male cells due to the process of X chromosome inactivation (XCI) [56]. Consequently, they are exposed to higher Prdm14 levels for a longer period, which could lead to the establishment of female-specific epigenetic marks. In fact, we previously reported that a Prdm14-responsive enhancer exhibited higher activity in female ES cells, strongly suggesting that Prdm14 target gene levels are dosage-sensitive [24]. In addition, it is possible that a subset of genes regulated by Prdm14 is distinct in male and female ES cells. This is also true of any dose-dependent TF or ERE with sex-biased expression. Therefore, future ChIP-seq studies for TFs and chromatin modifications performed in a sex-stratified manner should allow us to distinguish between these possibilities.

Our studies also show that there are factors in addition to Prdm14 that regulate sex-biased gene expression. For example, X-linked genes that are expressed from the two active X chromosomes, such as Atrx, Kdm6a, and Klf8, are strong candidates for involvement in sex-biased expression. However, autosomal factors, such as Lef1 and Zeb1, could be involved as well. In theory, TFs that are not sex-biased could also be important for differential gene expression and according to our network analysis (WGCNA), there are a host of other TFs and EREs that are significantly correlated with sex such as Arid3b, Smad4, Jarid2, and Kdm8. For regulatory factors that are not sex-biased per se, their cognate sites could present different accessibility in male and

---

Table 2 Female-biased expression of transcription and epigenetic factors (Continued)

| Transcription factors and epigenetic remodelling enzymes | ES | CP | 8.5 | 9.5 | 10.5 | p1 | Adult | Chr | Name |
|---------------------------------------------------------|----|----|-----|-----|------|----|-------|-----|------|
| Rsl1                                                    | XX | 13 | Regulator of sex limited protein 1 |
| Zbtb45                                                  | XX | 7  | Zinc finger and BTB domain containing 45 |
| Zfp282                                                  | XX | 6  | Zinc finger protein 282 |
| Zfp472                                                  | XX | 17 | Zinc finger protein 472 |
| Zfp498                                                  | XX | 5  | Zinc finger and SCAN domain containing 25 |
| Zfp758                                                  | XX | 17 | Zinc finger protein 758 |
| Zkscan14                                                | XX | 5  | Zinc finger with KRAB and SCAN domains 14 |
| Zkscan6                                                 | XX | 11 | Zinc finger with KRAB and SCAN domains 6 |
| Nat10                                                   | XX | 2  | N-acetyltransferase 10 |
| Senp3                                                   | XX | 11 | SUMO/sentrin specific peptidase 3 |
| Setmar                                                  | XX | 6  | SET domain without mariner transposase fusion |
| Sirt7                                                   | XX | 11 | sirtuin 7 |
| Suv39h2                                                 | XX | 2  | suppressor of variegation 3-9 2 |
| Bard1                                                   | XX | 1  | BRCA1 associated RING domain 1 |
| Cbx2                                                   | XX | 11 | Chromobox 2 |
| Cbx7                                                   | XX | 15 | Chromobox 7 |
| Rcc1                                                    | XX | 4  | Regulator of chromosome condensation 1 |
| Nko2-5                                                  | XX | 17 | NK2 homobox 5 |
| Irf4                                                    | XX | 13 | Interferon regulatory factor 4 |
| Cecr2                                                   | XX | 6  | CECR2, histone acetyl-lysine reader |

ES embryonic stem cells; CP cardiac precursors; 8.5, 9.5, 10.5 days post coitum (dpc) embryonic hearts (single cell); p1 neonatal hearts; Chr chromosome; empty cells expression not detected; nb expressed but not biased
| Transcription factors and epigenetic and remodelling enzymes | ES | CP | 8.5 | 9.5 | 10.5 | p1 | Adult | Chr | Name |
|---------------------------------------------------------------|----|----|-----|-----|------|----|-------|-----|------|
| Egr4                                                          | XY |     |     |     |      |    | Early growth response 4 |
| Cdx1                                                          | XY |     |     |     |      |    | Caudal type homeobox 1 |
| E2f7                                                          | XY |     |     |     |      |    | E2F transcription factor 7 |
| Eomes                                                          | XY |     |     |     |      |    | Eomesodermin |
| Evx1                                                          | XY |     |     |     |      |    | Even-skipped homeobox 1 |
| Fox3                                                          | XY |     |     |     |      |    | Forkhead box 3 |
| Foxp4                                                          | XY |     |     |     |      |    | Forkhead box P4 |
| Insr1                                                         | XY |     |     |     |      |    | Insulinoma-associated 1 |
| Lin28a                                                         | XY |     |     |     |      |    | Lin-28 homolog A (C. elegans) |
| Mesp1                                                         | XY |     |     |     |      |    | Mesenderm posterior 1 |
| Mbx1                                                          | XY |     |     |     |      |    | Mix1 homeobox-like 1 (Xenopus laevis) |
| Nr6a1                                                         | XY |     |     |     |      |    | Nuclear receptor subfamily 5, group A, member 1 |
| Pou2f3                                                        | XY |     |     |     |      |    | POU domain, class 2, transcription factor 3 |
| Sox11                                                         | XY |     |     |     |      |    | SRY (sex determining region Y)-box 11 |
| Sp5                                                           | XY |     |     |     |      |    | Per-hexamer repeat gene 4 |
| Sp8                                                           | XY |     |     |     |      |    | Trans-acting transcription factor 8 |
| T                                                             | XY |     |     |     |      |    | Brachyury, T-box transcription factor T |
| Tcf7                                                          | XY |     |     |     |      |    | Transcription factor 7, T cell specific |
| Wiz                                                           | XY |     |     |     |      |    | Widely-interspaced zinc finger motifs |
| Ybx2                                                          | XY |     |     |     |      |    | Y box protein 2 |
| Zglp1                                                         | XY |     |     |     |      |    | Zinc finger, GATA-like protein 1 |
| Dnmt3b                                                        | XY |     |     |     |      |    | DNA methyltransferase 3B |
| Dnmt3l                                                        | XY |     |     |     |      |    | DNA (cytosine-5’)-methyltransferase 3-like |
| Phc1                                                          | XY |     |     |     |      |    | Polyhomeotic 1 |
| Plac8                                                         | XY |     |     |     |      |    | Placenta-specific 8 |
| Tdrd5                                                         | XY |     |     |     |      |    | Tudor domain containing 5 |
| Bahcc1                                                        | XY |     |     |     |      |    | BAH domain and coiled-coil containing 1 |
| Kdm6c                                                         | XY | XY | XY | XY | XY | Y | Ubiq.transcribed tetraicopeptide repeat gene, Y chr |
| Bcl6b                                                         | XY | XY |     |     |     | nb | B cell CLL/lymphoma 6, member B |
| Irf8                                                          | XY | XY |     |     |     | nb | Interferon regulatory factor 8 |
| Atrid1a                                                       | XY | XY |     |     |     |    | AT rich interactive domain 1A (SWI-like) |
| Lef1                                                          | XY | XY | XY | XY | XY | 3 | Lymphoid enhancer binding factor 1 |
| Tbx20                                                         | XY | XY | XY | XY | XY | 9 | T-box 20 |
| Smarccl                                                       | XY | XY |     |     |     |    | SWI/SNF related, matrix assoc., actin-dep. reg. of chromatin d1 |
| Zbtb7a                                                        | XY | XY |     |     |     |    | Zinc finger and BTB domain containing 7a |
| Hif3a                                                         | XY | XY |     |     |     |    | hypoxia inducible factor 3, alpha subunit |
| Nfkb2                                                         | XY | XY | XY | XY | XY | 19 | nuclear factor of k light polypeptide enhancer B cells 2, p49/p100 |
| Pbx2                                                          | XY | XY |     |     |     | nb | Pre B cell leukemia homeobox 2 |
| Prdm6                                                         | XY | XY |     |     |     |    | PR domain containing 6 |
| Id2                                                           | XY | XY |     |     |     | 12 | Inhibitor of DNA binding 2 |
| Dot1l                                                         | XY | XY | XX |     |     | 10 | DOT1-like, histone H3 methyltransferase (S. cerevisiae) |
| Zfp296                                                        | XY | XX |     |     |     | 7 | Zinc finger protein 296 |
female cells or there could be differential availability of their co-factors.

**Differentiated ES cells exhibit sex-biased gene expression**

Differences in male and female ES cell lines into cardiac precursors drastically changed the transcriptional profile of the cells, but we still detected sex-biased expression. Most X-linked genes were expressed equally between male and female cells due to the process of XCI, but unexpectedly, four were more highly expressed in male cells, suggesting that there is male-specific regulation of some X-linked genes. Some of the sex differences in gene expression could represent the slight developmental delay of the female cells. Yet some expression differences observed in ES cells persist in the adult heart, suggesting that these are independent of developmental stage and are integrating *bona fide* sex-specific regulatory networks.

While the protocol we used for differentiation of ES cells into cardiac precursors has been derived from the extensive knowledge on cardiogenesis in vivo [41], the in vitro derivation of cardiac progenitors lacks other factors, such as spatial context, that are important for proper organ formation. For example, during heart development in vivo, multiple cell types, including transient populations, interact in three dimensions and receive input from surrounding tissues. However, single-cell analyses of early cardiac stages have pinpointed that cardiac progenitors derived from ES cells have a transcriptome corresponding to 9.5 dpc single-cell cardiomyocytes [43], a stage in which fibroblasts are not yet apparent. Thus, differentiated ES cells serve as a close approximation of the early stages of heart development.

**Sex-biased gene expression exists at every stage during cardiac development**

To determine whether the sex biases in differentiated ES cells are present in vivo, we inspected previously published data from specific stages during heart development. Single-cell assessment of transcriptional profiles in early stages of cardiogenesis has allowed detailed analysis of the step-wise specification of cardiac progenitors, but the available data are not stratified by sex. For each sample, we genotyped for sex and re-analyzed these data and observed sex-biased expression across all the available stages of heart development.

### Table 3 Male-biased expression of transcription and epigenetic factors (Continued)

| Transcription factors and epigenetic and remodelling enzymes | ES | CP | 8.5 | 9.5 | 10.5 | p1 | Adult | Chr | Name |
|-------------------------------------------------------------|----|----|-----|-----|------|----|-------|-----|------|
| Gata4                                                       | XY | nb | 14  |     |      |    |       |     | GATA binding protein 4 |
| Mycn                                                        | XY | nb | 12  |     |      |    |       |     | v-myc avian myelocytomatosis related og, neuroblastoma derived |
| Hdac5                                                       | XY | nb | 11  |     |      |    |       |     | Histone deacetylase 5 |
| Nfx1                                                        | XY | nb | 5   |     |      |    |       |     | Nuclear transcription factor, X-box binding-like 1 |
| Pax6                                                        | XY | nb | 2   |     |      |    |       |     | Paired box 6 |
| Brd4                                                        | XY | nb | 17  |     |      |    |       |     | Bromodomain containing 4 |
| Cargsp1                                                     | nb | XY | XY  | XY  | 16   |    |       |     | Calcium regulated heat stable protein 1 |
| Csde1                                                       | nb | XY | XY  | nb  | 3    |    |       |     | Cold shock domain containing E1, RNA binding |
| Mef2d                                                       | nb | XY | 3   |     |      |    |       |     | Myocyte enhancer factor 2D |
| Atf5                                                        | nb | XY | 7   |     |      |    |       |     | activating transcription factor 5 |
| Bach2                                                       | nb | XY | 4   |     |      |    |       |     | BTF and CNC homology, basic leucine zipper transcription factor 2 |
| Cdc5l                                                       | nb | XY | 17  |     |      |    |       |     | Cell division cycle 5-like (S. pombe) |
| Crebf                                                       | nb | XY | 7   |     |      |    |       |     | CREB/ATF bZIP transcription factor |
| Gm13139                                                     | nb | XY | 4   |     |      |    |       |     | Zinc finger protein 991 |
| Hmgb1                                                       | nb | XY | 2   |     |      |    |       |     | Predicted gene, 21596 |
| Hmgb2                                                       | nb | XY | 8   |     |      |    |       |     | High mobility group box 2 |
| Klf10                                                       | nb | XY | 15  |     |      |    |       |     | Kruppel-like factor 10 |
| Nfic                                                        | nb | XY | 10  |     |      |    |       |     | Nuclear factor Ic |
| Nkb1                                                        | nb | XY | 3   |     |      |    |       |     | Nuclear factor of κ light polypeptide enhancer in B cells 1, p105 |
| Pias2                                                       | nb | XY | 18  |     |      |    |       |     | Protein inhibitor of activated STAT 2 |

**ES** embryonic stem cells; **CP** cardiac precursors; **8.5, 9.5, 10.5 days post coitum (dpc) embryonic hearts (single cell); **p1** neonatal hearts; **Chr** chromosome; **empty cells** expression not detected; **nb** expressed but not biased
development. We also observed short bursts of sex-biased expression of regulatory factors at single stages, raising the question of whether these are capable of encoding persistent dimorphisms. In addition, we show that some genes equalize their expression, while others become biased in the opposite direction, which raises important questions on the mechanisms by which these events occur.
We recognize several caveats in this study. First, compiling datasets from different reports presents challenges because of the different experimental designs. Our own data is from ES cells in culture subjected to a directed differentiation protocol that only partially recapitulates the complex processes in vivo. Second, the single-cell RNA-seq data from embryonic and neonatal hearts, while useful for distinguishing cell populations, is necessarily incomplete. Currently, single-cell RNA-seq only detects a fraction of the transcriptome, with a bias towards high expression transcripts, which excludes many TFs that are expressed at relatively low levels.

Systems-level analyses have yielded valuable information on the correlations between congenital heart disease and their developmental origins [62, 63]. Transcriptomic data for early developmental stages is sparse, however. Nevertheless, the currently available datasets reveal sex-biased expression at every stage and suggest novel hypotheses for future mechanistic studies. Our analyses also open questions on how the fluctuations in sex-biased expression are regulated, how they are reflected in epigenetic differences between male and female cells, and how widely these occur in other tissues during embryogenesis. Our data also serves as a platform to identify the role of sex hormones in counteracting or compounding sex biases. Future studies will enable dissection of the effects of sex chromosomes and hormonal influence on sexual dimorphism. Ultimately, expanding developmental studies will allow us to connect early sexual dimorphism to the sex biases that occur in adult health and disease.

Conclusions

The ability to profile transcriptomes has heightened interest in sex-biased gene expression, especially after recent reports that show substantial differences between males and females in humans and other animal models, even in organ systems that are overtly identical [4, 5, 44, 64]. The focus on adult tissues reflects a broadly held assumption that sex-biased expression is unimportant during early embryogenesis, during which critical lineage decisions are made, and that sex-specific selection only operates after the reproductive interests of the sexes have diverged [65]. In non-mammalian species, however, there is evidence that sex biases at the transcriptomic level occur throughout development [17, 66]. Here, we address a major gap in developmental studies by detecting sex-biased expression during mouse cardiac development. Our data strongly suggest that some of the differences in transcriptomic profiles in adult hearts may be established epigenetically before the appearance of sex hormones. Our observations open the field to explore the timing and extent of sex-specific transcriptional and epigenetic profiles in other organ systems and their relevance to sexual dimorphisms in adult health and disease.

Additional files

**Additional file 1:** Table S1. List of antibodies used for ChIP-seq (XLSX 9 kb)
**Additional file 2:** Table S2. Accession and metadata of all samples analyzed in this study (XLSX 10 kb)
**Additional file 3:** Dataset S1. Genes in blue/violet module eigengene (XLSX 29 kb)
**Additional file 4:** Dataset S2. Ingenuity Pathway Analysis Upstream Regulators (XLSX 9 kb)
**Additional file 5:** Dataset S3. Topology analysis for protein-protein interaction networks (XLSX 141 kb)
**Additional file 6:** Figure S1. Modularity in male and female ES cell protein-protein interaction networks (JPG 50 kb)
**Additional file 7:** Dataset S4. Gene ontology for protein-protein interaction networks (XLSX 15 kb)
**Additional file 8:** Figure S2. qPCR of markers before and after ES cell differentiation (JPG 38 kb)
**Additional file 9:** Dataset S5. Differentially expressed genes in male and female cardiac precursors (XLSX 10 kb)
**Additional file 10:** Figure S3. Male-enriched genes after differentiation of ES cells. UCSC browser screenshot with tracks denoting Prdm14 occupancy (track obtained from Ma et al.) (JPG 41 kb)
**Additional file 11:** Dataset S6. Sex Biases in Single Cell Data (XLSX 108 kb)
**Additional file 12:** Figure S4. Modularity in protein-protein interaction networks (JPG 114 kb)
**Additional file 13:** Figure S5. Differentially expressed genes in adult male and female hearts (JPG 113 kb)
**Additional file 14:** Datasets S7 and S8. Ingenuity Pathway Analysis for sex-biased genes in adult heart (XLSX 19 kb)
**Additional file 15:** Dataset S9. Sex-biased Genes Shared between Human and Mouse Adult Hearts (XLSX 13 kb)

**Abbreviations**

EREs: Epigenetic and remodeling enzymes; ES cells: Embryonic stem cells; FPKM: Fragments per kilobase of exon per million reads mapped; GS: Gene significance; LIF: Leukemia inhibitory factor; ME: Module eigengene; MS: Module significance; PPis: Protein-protein interaction networks; TFs: Transcription factors; TO: Topology overlap; WGCNA: Weighted gene co-expression network analysis; XCI: X chromosome inactivation

**Acknowledgments**

We thank Rachael Werner for technical support and Joanne Thorvaldsen for comments and suggestions. This project was supported by a Temple University Bridge Grant (to N.E.).

**Authors’ contributions**

NE conceived and designed the experiments. DFD generated the data. RK and RJK performed the data analysis. NE and RJK interpreted the data. JM contributed bioinformatic support. NE wrote the manuscript. NE, JM and RK generated the figures. All authors gave final approval of the published version. All authors are accountable for all aspects of the work.

**Funding**

This study was funded by a Temple University Bridge Grant. This research includes calculations carried out on Temple University’s HPC resources and thus was supported in part by the National Science Foundation through
major research instrumentation grant number 1625061 and by the US Army Research Laboratory under contract number W911NF-16-2-0189.

Availability of data and materials
Data generated has been deposited in GEO: GSE90516. Data from other reports and their supplementary information files was also analyzed.

References
1. Blenck CL, Harvey PA, Reckelhoff JF, Leinwand LA. The importance of cardiac gene expression during heart development. Dev Cell 2016, 39(4):491-507. GSE76118
2. DeLaughter DM, Bick AG, Wakimoto H, McKeen D, Gorham JM, Kathiria IS, Hinson JT, Hosmy J, Gray J, Pu W et al. Single-Cell Resolution of Temporal Gene Expression during Heart Development. Dev Cell 2016, 39(4):480-490. Obtained from the author.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1Fels Institute for Cancer Research, Lewis Katz School of Medicine, Temple University, 3400 N. Broad St, Philadelphia, PA 19140, USA. 2Department of Biology, College of Science and Technology, Temple University, 1900 N. 12th St, Philadelphia, PA 19122, USA.

Received: 12 June 2019 Accepted: 20 August 2019

References
1. Blenck CL, Harvey PA, Reckelhoff JF, Leinwand LA. The importance of biological sex and estrogen in rodent models of cardiovascular health and disease. Circ Res. 2016;118(8):1294–312.
2. Singmann P, Shem-Tov D, Wahl S, Graffit H, Fierro G, Shin SY, Schramm K, Wolf P, Kunze S, Baran Y et al. Characterization of whole-genome autosomal differences of DNA methylation between men and women. Epigenetics Chromatin. 2015;8:43.
3. Isensee J, Ruiz Noppinger P. Sexually dimorphic gene expression in mammalian somatic tissue. Gender medicine. 2007;4(Suppl B):S75–S79.
4. Mank JE, Nam K, Brunstrom B, Ellegren H: ontogenetic complexity of sexual dimorphism and sex-specific selection. Mol Biol Evol. 2010;27(7):1570–8.
5. Durin SJ, Miartello G, Yordanov B, Emmott S, Smith AG. Defining an essential transcription factor program for male pluripotency. Science. 2014;346(6211):155–60.
6. Gershoni M, Pietrokovski S. The landscape of sex-differential transcriptome expression and regulatory networks across human tissues. bioRxiv 2016.
7. Isensee J, Witt H, Pregla R, Hetzer R, Regitz-Zagrosek V, Noppinger PR. Sexually dimorphic transcriptional program of mouse embryonic stem cells. Stem Cells. 2017;35(8):2535–43.
8. Yu Y, Fuscoe JC, Zhao C, Guo C, Jia M, Qing T, Bannon DI, Lancashire L, Bao LW, Du T, et al. A rat RNA-Seq transcriptomic BodyMap across 11 organs and its consequent selection in human adults. BMC Biol. 2017;15(1):7.
9. Jirous Z, Locci M, Mancini M. The Denolin lecture: the woman with congenital heart disease. Eur Heart J. 1998;19(12):1766–75.
10. Jenkins KJ, Correa A, Feinstein JA, Botto L, Britt AE, Daniels SR, Elinson M, Waters CA, Webb CL. American Heart Association Council on Cardiovascular Disease in the Y: Noninherited risk factors and congenital cardiovascular defects: current knowledge, a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation. 2007;115(23):2995–3014.
11. Deegan JD, Moskowitz IP. TBX5: A key regulator of heart development. Curr Top Dev Biol. 2017;122:195–221.
12. Bardot E, Calderon D, Santoriello F, Han S, Cheung K, Jadhav B, Burtscher I, Artap S, Jain R, Epstein J et al. Foxa2 identifies a cardiac progenitor population with ventricular differentiation potential. Nat Commun. 2017;8:14428.
13. Azhar M, Ware SM. Genetic and developmental basis of cardiovascular malformations. Clin Perinatol. 2016;43(1):39–53.
14. Somerville J. The Denolin lecture: the woman with congenital heart disease. Eur Heart J. 1998;19(12):1766–75.
15. Jenkins KJ, Correa A, Feinstein JA, Botto L, Britt AE, Daniels SR, Elinson M, Waters CA, Webb CL. American Heart Association Council on Cardiovascular Disease in the Y: Noninherited risk factors and congenital cardiovascular defects: current knowledge, a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation. 2007;115(23):2995–3014.
16. Mank JE, Nam K, Brunstrom B, Ellegren H: ontogenetic complexity of sexual dimorphism and sex-specific selection. Mol Biol Evol. 2010;27(7):1570–8.
17. Durin SJ, Miartello G, Yordanov B, Emmott S, Smith AG. Defining an essential transcription factor program for male pluripotency. Science. 2014;346(6211):155–60.
18. Gershoni M, Pietrokovski S. The landscape of sex-differential transcriptome expression and regulatory networks across human tissues. bioRxiv 2016.
19. Jirous Z, Locci M, Mancini M. The Denolin lecture: the woman with congenital heart disease. Eur Heart J. 1998;19(12):1766–75.
20. Deegan JD, Moskowitz IP. TBX5: A key regulator of heart development. Curr Top Dev Biol. 2017;122:195–221.
