

## ARTICLE

# Human active X-specific DNA methylation events showing stability across time and tissues

Jihoon Eric Joo<sup>1</sup>, Boris Novakovic<sup>1,2</sup>, Mark Cruickshank<sup>3</sup>, Lex W Doyle<sup>2,4,5</sup>, Jeffrey M Craig<sup>\*,2,6</sup>  
 and Richard Saffery<sup>1,2</sup>

The phenomenon of X chromosome inactivation in female mammals is well characterised and remains the archetypal example of dosage compensation via monoallelic expression. The temporal series of events that culminates in inactive X-specific gene silencing by DNA methylation has revealed a ‘patchwork’ of gene inactivation along the chromosome, with approximately 15% of genes escaping. Such genes are therefore potentially subject to sex-specific imbalance between males and females. Aside from *XIST*, the non-coding RNA on the X chromosome destined to be inactivated, very little is known about the extent of loci that may be selectively silenced on the active X chromosome (Xa). Using longitudinal array-based DNA methylation profiling of two human tissues, we have identified specific and widespread active X-specific DNA methylation showing stability over time and across tissues of disparate origin. Our panel of X-chromosome loci subject to methylation on Xa reflects a potentially novel mechanism for controlling female-specific X inactivation and sex-specific dimorphisms in humans. Further work is needed to investigate these phenomena.

*European Journal of Human Genetics* (2014) **22**, 1376–1381; doi:10.1038/ejhg.2014.34; published online 9 April 2014

## INTRODUCTION

The potential for a discrepancy in gene activity because of X chromosome dosage differences between male and female mammals is largely overcome during early development<sup>1</sup> by the inactivation of one of two X chromosomes in females.<sup>2</sup> Based largely on studies in mice, it appears that the paternally derived copy of the X chromosome (Xp) is preferentially inactivated by imprinting during early embryogenesis,<sup>3</sup> but then reactivated at the blastocyst stage of development (imprinted X chromosome inactivation (XCI)). Placental mammals then undergo a second wave of XCI, which results in the silencing of most genes on a randomly selected copy (random XCI) of the X chromosome (inactive X (Xi); reviewed in Morey and Avner<sup>4</sup>).

The X-inactivation centre (XIC) at Xq13 regulates both phases of XCI.<sup>5</sup> This region contains important genes for XCI, including *XIST* (the Xi specific transcript<sup>6</sup>), a non-coding RNA that has a key role in silencing Xi.<sup>6–8</sup> *XIST* is expressed exclusively from the chromosome destined for inactivation and works *in cis* to modulate epigenetic silencing of the chromosome.<sup>9</sup> The promoter of *XIST* on the active copy of the X chromosome (Xa) is silenced and methylated, whereas the *XIST* promoter on Xi is hypomethylated.<sup>10</sup>

The process of XCI is incomplete, with approximately 15% of genes on human X chromosomes ‘escaping’ inactivation, potentially leading to gene expression imbalance between sexes.<sup>11</sup> Gene expression and DNA methylation studies in humans have identified more than a dozen similar genes escaping X inactivation.<sup>12,13</sup> However, *XIST* remains the only gene demonstrated to be methylated and silenced specifically on Xa. Nevertheless, recent DNA methylation profiling of the human active and Xi chromosome using an affinity-based

approach revealed multiple regions of apparently lower levels of methylation on the X chromosome in XX females relative to XO females, suggesting an association with active X-specific methylation.<sup>14</sup> In the current study, we used high-resolution DNA methylation BeadChip arrays to investigate the extent of Xa-specific methylation in tissues of different origin collected longitudinally from birth.

## MATERIALS AND METHODS

### Samples used in this study

Samples and the Infinium HM450 data set used in this study largely overlap with our published studies on preterm birth<sup>15</sup> and our longitudinal DNA methylation changes in twins.<sup>16</sup> Buccal samples used in this study were selected based on the availability of the materials and array data quality. Data from dried blood spots used in this study were from a case/control study of DNA methylation associated with preterm birth. No sample was excluded from this study other than those with poor quality of array data.

### DNA isolation and bisulfite conversion

Blood DNA used in this study was obtained from the Victorian Infant Collaborative Study Cohort.<sup>17</sup> Six to ten 3-mm dried blood spot punches per participant were digested in 200 µl of PBS (Life Technologies, Grand Island, NY, USA) and 20 µl proteinase K (Qiagen, Hilden, Germany), and homogenised using a Tissue Lyser (Qiagen). Genomic DNA was isolated using Qiagen Blood minispin kits, following the manufacturer’s instructions. Buccal cells were obtained from the Peri/postnatal Epigenetic Twins Study.<sup>18</sup> Buccal DNA was isolated by traditional salting out method as previously described.<sup>19</sup> Buccal cells were detached from swabs by agitation in PBS. The cells were digested by proteinase K overnight and the lysate was centrifuged. The supernatants containing DNA were transferred to fresh tubes and DNA was precipitated by 100% ethanol and 3 M sodium acetate. DNA was pelleted

<sup>1</sup>Cancer and Disease Epigenetics, Murdoch Childrens Research Institute, Royal Children’s Hospital, Parkville, Victoria, Australia; <sup>2</sup>Department of Paediatrics, University of Melbourne, Parkville, Victoria, Australia; <sup>3</sup>Leukaemia and Cancer Research, Telethon Institute for Child Health Research, University of Western Australia, Perth, Western Australia, Australia; <sup>4</sup>Neonatal Services, Royal Women’s Hospital, Parkville, Victoria, Australia; <sup>5</sup>Department of Obstetrics and Gynaecology, University of Melbourne, Royal Women’s Hospital, Parkville, Victoria, Australia; <sup>6</sup>Early Life Epigenetics, Murdoch Childrens Research Institute, Royal Children’s Hospital, Parkville, Victoria, Australia

\*Correspondence: Dr JM Craig, Early Life Epigenetics, Murdoch Childrens Research Institute, Royal Children’s Hospital, Flemington Road, Parkville, Victoria 3052, Australia. Tel: +61 3 8341 6346; Fax: +613 9348 1391; E-mail: jeff.craig@mcri.edu.au

Received 3 June 2013; revised 29 January 2014; accepted 13 February 2014; published online 9 April 2014

by centrifugation and resuspended in TE buffer. Placental samples were also obtained from the Peri/postnatal Epigenetic Twins Study.<sup>18</sup> Multiple X cell lines used in the study were patient de-identified skin fibroblasts with 48,XXXX and 47,XXX karyotypes. A measure of 500 ng of genomic DNA was subjected for bisulfite conversion using MethylEasy Exceed (Human Genetics Signature, Sydney, Australia). The conversion efficiency was assessed by bisulfite-specific PCR, before hybridising to the HM450 BeadChip (Illumina, Scoresby, VIC, Australia).

All samples in this study were obtained with written informed consents from all participants (parents and subjects). This study was approved by the Human Research Ethics Committees, Royal Children's Hospital and Royal Women's Hospital (Melbourne, Australia).

### Illumina infinium human methylation HM450

Hybridization of bisulfite-treated samples to Illumina Infinium Human Methylation450 (HM450) BeadChips was performed at the Service XS (Leiden, The Netherlands). Raw signals were exported to R statistical environment (<http://cran.r-project.org/index.html>, v2.15.1) and processed using the *minfi* package (v1.3.15, Release date 24 September 2012) available from Bioconductor.<sup>20</sup> To minimise the discrepancy between type I and type II probes, the data were normalised using Subset-quantile Within Array Normalisation.<sup>21</sup> All X chromosome probes were then extracted according to the HumanMethylation450 annotation file v 1.2, which was obtained from Illumina (<http://www.illumina.com>). Poor performing probes were removed using a detection *P*-value cutoff, 0.02 for all samples.  $\beta$ - and *M*-values were calculated using the *minfi* package. *B*-values were calculated by following formula: Methylated probes/(Methylated + Unmethylated probes + 100) and *M*-values were calculated by:  $\log_2 \{(\text{methylated probes} + 1)/(\text{unmethylated probes} + 1)\}$ . Generally speaking,  $\beta$  and *M* values have a base 2 logistic relationship.<sup>19</sup>

Linear regression analyses were performed on *M*-values using the *limma* package (v3.14.0, Release date 25 September 2012).<sup>22</sup> Adjusted *P*-values were calculated using False Discovery Rate control method integrated in the *limma*. All DNA methylation calculations were also carried out in R.

### Sequenom EpiTYPER

Locus-specific DNA methylation analysis to validate our findings was performed using Sequenom EpiTYPER MassARRAY (Sequenom, San Diego, CA, USA), as previously described.<sup>19</sup> Bisulfite primers were designed using EpiDesigner (<http://www.epidesigner.com>, version number is not available). Details of primers used for EpiTYPER analysis can be found in Supplementary Figure 1. Bisulfite-converted DNA was amplified using FastStart PCR Mastermix (Roche, Penzberg, Germany). Sequenom EpiTYPER MassARRAY was set up according to the manufacturer's instruction. T-cleavage enzyme was used for fragmentation of amplicons. All PCR and EpiTYPER were performed in technical triplicates and mean methylation scores were used for analysis.

### Gene expression dataset

A raw gene expression data set of monocytes and T lymphocytes isolated from 20 healthy individuals (10 males/10 females) was acquired from The Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>; Accession ID: GSE30059) and processed using the GEOquery R package (v2.62.2, release date 07 April 2013).<sup>23</sup> The annotation for Illumina Human HT-12 V3.0 expression BeadChip (Platform ID: GPL6947) was also downloaded using same package. To identify differentially expressed genes, a linear regression analysis was performed using the *limma* R package.<sup>22</sup> Adjusted *P*-values were calculated using FDR method available from the *limma*.

## RESULTS AND DISCUSSION

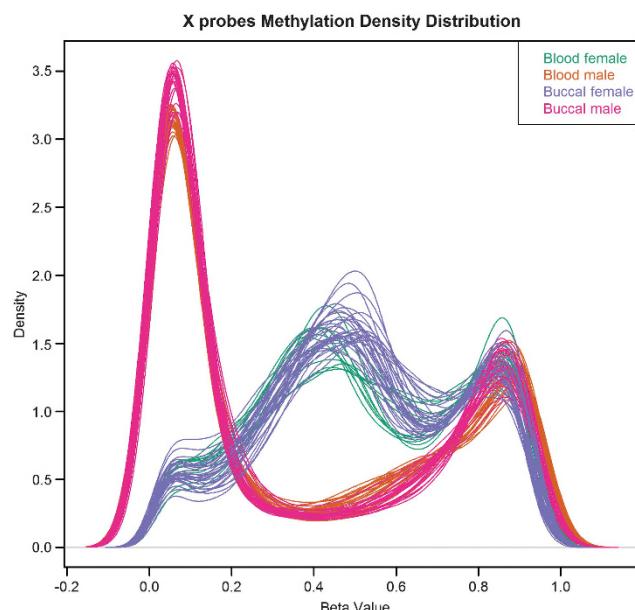
### Distribution of X chromosome DNA methylation

The Infinium HumanMethylation450 BeadChip platform (HM450) interrogates the methylation status of >480 000 CpG dinucleotides across the genome, including 11 232 on the X chromosome. Data from these sex-specific probes are often excluded from epigenome-wide association studies because sex effects represent one of the largest confounders in comparative genome-scale DNA methylation studies.

We generated longitudinal DNA methylation data from 38 whole-blood spots (including 26 male and 12 female samples) sampled at birth (14) and 18 years (24 samples) from singletons and from 56 buccal cell samples (25 males and 31 females) sampled at birth (31 samples) and 18 months (25 samples) of age from monozygotic and dizygotic twin pairs (Supplementary Table 1). Twenty-eight of the blood samples were collected from the same individuals at birth and 18 years old (ie, 14 pairs). For buccal cells, both siblings of twin pairs were used in the analysis such that the data set comprised a total of 42 longitudinally matched twin buccal samples (ie, 21 pairs). Sibling samples were treated as different individuals unless stated otherwise.

Average detection *P*-values of all detected probes were less than 0.02 for all 94 samples, indicating high-quality data. For all, but two, of these samples, average detection *P*-values were less than 0.01. Probes with detection *P*-values above 0.05 in any sample, indicative of potentially unreliable methylation calls, were removed from the entire data set, leaving a total of 9558 X-chromosome probes common across all samples. Principal components analysis on all detected probes revealed tissue-specific methylation as the major source of variation with no apparent outlying samples in either tissue group (Supplementary Figure 2).

The  $\beta$ -value density distribution of X-chromosome probes revealed distinctly sex-specific DNA methylation patterns irrespective of the tissue type or time point (Figure 1). Most probes in females (>60%; 6016 in buccal, 5742 in blood) showed a pattern consistent with intermediate methylation ( $\beta$ -values between 0.2 and 0.7). We also defined sets: hypomethylated ( $\beta$ -values <0.2) and hypermethylated ( $\beta$ -values >0.7) probes, which together reveal an overall trimodal distribution of methylation peaks in females. The majority of these probes (4352 in buccal, 4073 in blood) were hypomethylated in males ( $\beta$ -values <0.2; Figure 1), and these showed an overall bimodal distribution (Figure 1). This represents a typical distribution pattern previously associated with the Xi chromosome.<sup>1</sup>



**Figure 1** X chromosome DNA methylation density plot.  $\beta$ -Value density plot showing DNA methylation distribution of X chromosome probes in males and females in blood and buccal DNA. The majority of probes in males are consistently hypomethylated ( $\beta$  values <0.2) or hypermethylated ( $\beta$  values >0.7), whereas a large proportion of probes (40–50%) are hemi-methylated (0.2 <  $\beta$  < 0.7) in females.

### Identification of constitutive Xa-associated DNA methylation

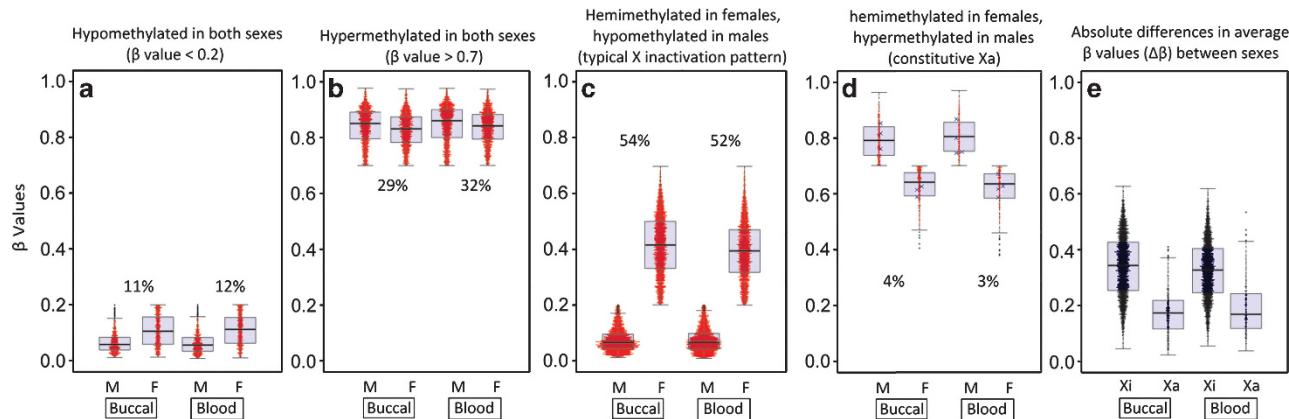
X chromosome methylation levels for each probe in the combined data set were then classed into one of three groups according to average methylation level across all samples of both sexes combined similar to that described previously.<sup>13</sup> Hypomethylated probes were defined as average  $\beta$  value less than 0.2, hypermethylation as  $\beta > 0.7$  and intermediate methylation was classified as between  $0.2 < \beta < 0.7$ . As anticipated, the majority of probes (52–54% of the total) showed the conventional XCI methylation patterns, which we define as differentially methylated probes (DMPs; see below) that are hypomethylated in males relative to females in both tissues at both time points in each tissue (ie, birth and 18 years for blood; birth and 18 months for buccals; Figure 2c; Supplementary Tables 4 and 5). A total of 909 (11%; buccal) and 917 (12%; blood) probes were consistently hypomethylated in both sexes (Figure 2a) and 2369 (29%; buccal) and 2461 (32%; blood) probes were hypermethylated in both sexes (Figure 2b) at both time points in each tissue.

To identify X chromosome probes that were methylated in a sex-specific manner, we first performed linear regression analyses to identify statistically significant DMPs between males and females including those probes showing an XCI pattern, present at both time points for each tissue type. In buccal cell samples, 8226 DMPs were identified when birth and 18 months data were combined (adjusted  $P$ -values  $< 0.05$ , Supplementary Table 2). In blood, 7299 probes were found to be significantly different between sexes at both birth and 18 years combined (Supplementary Table 3).

In addition to the expected Xi-specific methylation (DMPs that are hypomethylated in males and intermediately methylated in females, Figure 2c), approximately 3–4% of DMPs were found to be consistently hypermethylated in males, with intermediate methylation levels in females, in both tissues and at both time points (Figure 2d). This is consistent with Xa-specific methylation. Importantly, these probes included several probes at the *XIST* non-coding RNA locus, consistent with previous findings of Xa-specific methylation of this gene<sup>10</sup> (Supplementary Tables 4 and 5). However, the degree of absolute methylation differences ( $\Delta\beta$ ) between sexes was much

smaller in probes showing Xa-specific methylation (including the *XIST* probes) than those showing X inactivation methylation pattern (Figure 2e). At this stage, we do not have a clear explanation on this, although it could be an artefact because of the platform's uneven heteroscedastic  $\beta$  distribution. Nevertheless, the sex-specific differences in our Xa probes are much larger than between-sex variations in non-Xa probes (data not shown). A further subset of probes showed either tissue-specific or timing-specific Xa-associated methylation and were not investigated further.

To further validate our findings of Xa-specific methylation from the HM450 arrays and in another tissue-type, we carried out locus-specific methylation analysis of multiple CpGs sites located around transcription start sites of the *BCOR* gene in 12 placental samples collected at birth. Such cross-platform validation ensures that our finding is not technical artefacts and is reproducible using other detection methods. Furthermore, by using another tissue type, we tested whether this phenomenon is also seen in other tissue types. *BCOR* encodes the BLC6 co-repressor protein previously linked to certain cancers<sup>24,25</sup> and was selected as the gene overlapped 25 HM450 probes showing a distinct Xa-specific methylation pattern and had the most extensive probe coverage on the X chromosome after *XIST*. The placenta was chosen as it represents the extraembryonic lineage, whereas blood is mesodermal and buccal ectodermal in origin, enabling us to test whether this Xa-specific pattern in this region is limited to specific tissue types. Our HM450 array data showed that the mean differences in absolute  $\beta$ -values between males and females ranged from 0.11 to 0.53 in *BCOR* probes in both buccal tissues and blood (median = 0.26). Two regions of Xa-specific differential methylation, as identified by HM450 analysis, were selected for cross-platform and tissue validation by Sequenom MassArray EpiTYPER analysis. Methylation levels at 23 CpG dinucleotides, including 2 coinciding with HM450 probes, were measured in placental samples ( $n = 6$  males versus  $n = 5$  females). We observed Xa-specific methylation patterns across both regions, suggesting that Xa methylation at these regions of *BCOR* may be a constitutive state in humans (Figure 3). Further, we observed reduced



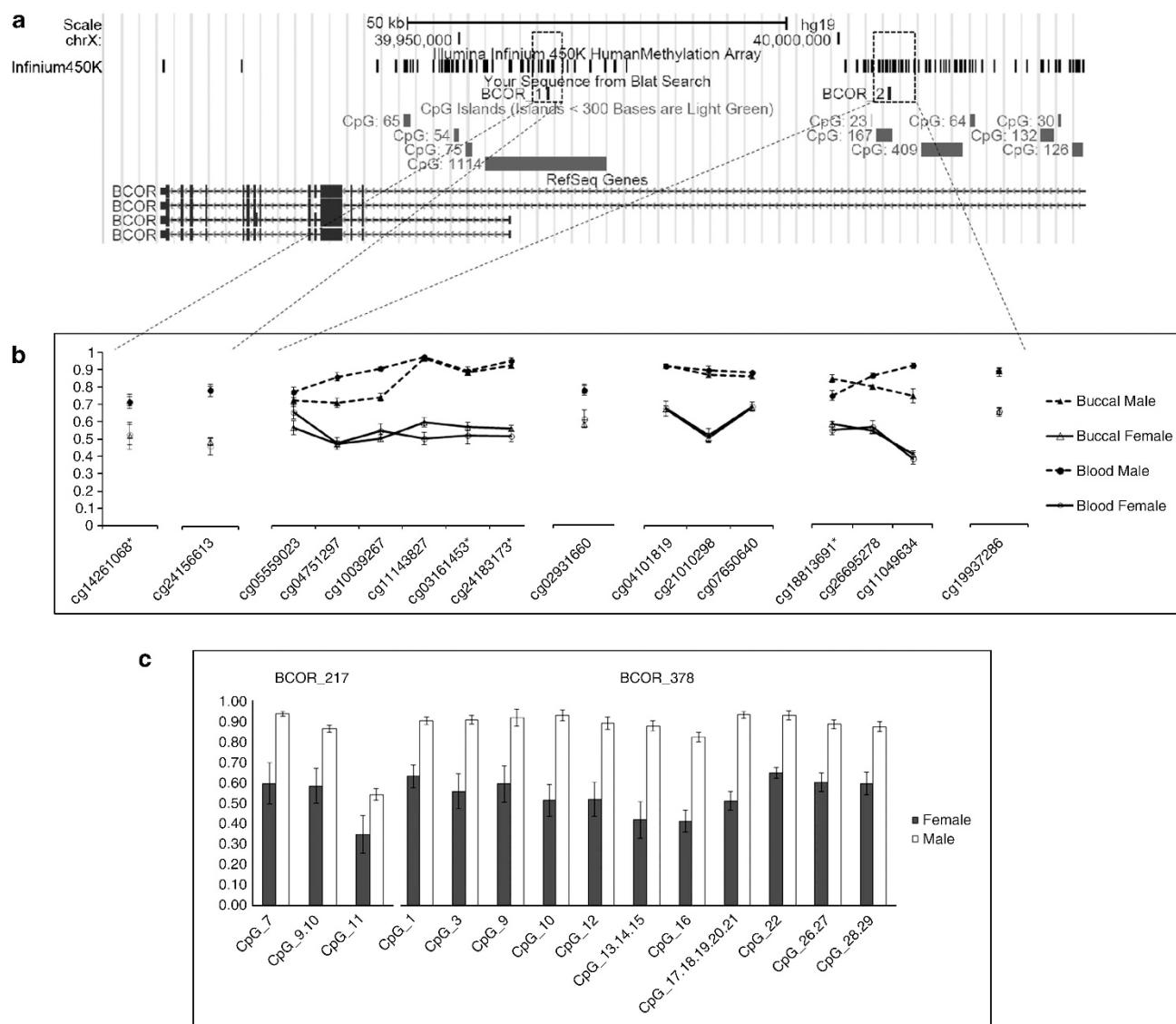
**Figure 2** Boxplots overlayed with one-dimensional scatterplots showing DNA methylation distributions for specific groups of probes. Average  $\beta$ -values were calculated for each probe group and separated by tissue type and sex (a). Hypomethylated probes ( $\beta$  values  $< 0.2$ ) in both sexes (b). hypermethylated probes ( $\beta$  values  $> 0.7$ ) in both sexes. Hypomethylated (shown in a) and hypermethylated (b) probes show no significant methylation differences between males and females and are speculated not to have any sex-specific biological roles. (c) Probes showing an X-inactivation methylation pattern (ie, hypomethylated in males and hemimethylated ( $\beta$  values  $> 0.2$  and  $< 0.7$ ) in females), CpGs in c show the typical XCI methylation pattern, where one of two copies of female X chromosomes is randomly inactivated by DNA methylation to dosage compensate the gene expression from single copy of X chromosome in males. (d) Probes showing an Xa-specific methylation pattern (ie, hypermethylated in males and hemimethylated in females). Xa-specific probes shown in d may escape the dosage compensation and have potential roles in controlling genes involved in sex-specific dimorphisms (eg, *XIST*).  $\beta$ -Values of the (Xa) *XIST* probes are shown as blue crosses. (e) Absolute between-sex differences in  $\beta$ -values ( $\Delta\beta$ ) of X inactivation (Xi) and constitutive Xa probes. Approximate % of probes in each category and sex is indicated.

levels of methylation of these regions in a small number of fibroblast cell lines derived from individuals with two copies of Xi and one of Xa (XXX) and three copies of Xi and one of Xa (XXXX) relative to cells with one Xa and one Xi (XX), consistent with Xa-associated methylation (Supplementary Figure 3).

### Xa-specific methylation is associated with genes

We further examined the genomic context of the Xa-specific methylation in males to identify associated genes and underlying sequence features. We identified 423 and 291 Xa-methylated probes, linked to 243 and 172 unique genes in buccal and blood, respectively (Supplementary Tables 4 and 5). By comparing the two tissue-specific Xa lists, we identified 170 probes in common, linked to 94 unique genes, which we designated 'Xa constitutive' in both tissues. These genes/probes were situated across the X chromosome and not skewed

towards any specific X chromosome regions, including the two major pseudoautosomal regions PAR1<sup>26</sup> and PAR2.<sup>27</sup> None of the Xa-specific methylation marks overlapped key pseudoautosomal genes.<sup>26,27</sup> Furthermore, we were unable to detect any specific gene ontologies overrepresented within our Xa constitutive genes (data not shown). Three Xa constitutive probes mapped to the *XIST* promoter as expected. A total of 84 (55%) Xa constitutive probes were promoter associated and/or located within 1500 bp upstream of a transcription start site (Table 1). In contrast, 70% of constitutively hypomethylated probes were promoter associated, whereas only 45% of constitutively hypermethylated probes were promoter associated (Table 1,  $\chi^2$  (expected/observed) test:  $P < 0.05$ ). This is consistent with a previous finding of hypermethylation of intragenic and intergenic regions on the X chromosome.<sup>13</sup> Furthermore, notably smaller proportions of hypermethylated probes were associated with enhancers ( $\sim 2\%$ ) and DNAse hypersensitive sites ( $\sim 1\%$ ), a



**Figure 3** Sex-specific DNA methylation at the *BCOR* promoter and CpG islands. (a) Genomic regions showing clear differential methylation patterns between males and females in both tissues and time points. BCOR\_1 and BCOR\_2 denote regions where the methylation was also assessed by Sequenom EpiTYPER MassARRAY. (b) Differential methylation patterns between males and females detected by the HM450, \*HM450 probe CpG sites that were also assessed by Sequenom EpiTYPER. (c) Differential methylation patterns of 12 placental tissue samples in two regions detected by EpiTYPER, the regions clearly show a sex-specific pattern. Error bars denote 95% confidence intervals.

**Table 1** Analysis of the HM450 X chromosome probes annotations

	All detected X probes (%)	Constitutive X inactivation (%)	Constitutive X activation (%)	Constitute hypomethylated (%)	Constitute hypermethylated (%)
Promoter	59	64	55	70	45
Non-promoter	41	36	45	30	55
Enhancer	6	5	6	5	2
Non-enhancer	94	95	94	95	98
DHS	18	32	2	30	1
Non-DHS	82	68	98	70	99
CGI (shelf)	9	<1	16	1	17
CGI (shore)	24	22	27	18	22
CGI	43	70	17	70	18
Non-CGI	27	8	40	11	45

Abbreviations: CGI, CpG island; DHS, DNase hypersensitive site.

Probes associated with promoters were selected by the location of probes in relation to transcription start sites (<1500 bp from TSS, 5'UTR). Relationship to enhancer and CGIs was determined using Illumina's annotation v 1.2. DHS (DNase hypersensitivity)-associated probes were determined according to the Encyclopaedia of DNA Elements (ENCODE) database (<http://www.genome.ucsc.edu/ENCODE/>). All detected X probes: 9558 X chromosome probes detected in all 99 samples. Constitutive X inactivation: probes consistently showing hypomethylation in males and intermediate methylation in females. Constitutive X activation: hypermethylation in males and intermediate methylation in females. Constitutive hypomethylated: hypomethylation in both males and females. Constitutive hypermethylated: hypermethylation in both males and females.

predictor for regulatory regions, when compared with all X chromosome probes (Table 1). In addition, Xa-specific hypermethylation was over-represented at CpG island (CGI) shelves (flanking regions of 2–4 kb from CGI), non-CGI regions (Open Sea regions >4 kb away from CGI;  $\chi^2$  test:  $P<0.05$ ), and were under-represented at CGIs and DNase hypersensitive site regions ( $\chi^2$  test:  $P<0.05$ ).

#### Xa-specific methylation does not necessarily correlate with gene expression

To further investigate whether gene expression changes are associated with our Xa-specific methylation marks, we assessed a publicly available gene expression data set of monocytes and T lymphocytes isolated from 20 healthy donors from the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>). These donors were aged between 23 and 45 years at the time of blood collection and consist of 10 females and 10 males. By fitting a linear regression model, we identified 12 transcripts from the X chromosome that were differentially expressed between males and females (FDR adj.  $P$ -value  $<0.05$ , Supplementary Table 6). Of these 12 transcripts, 10 were overexpressed in females (Supplementary Figure 4). Consistent with previous reports, the *XIST* gene showed the most highly significant sex-specific levels between two groups (FDR adj.  $P$ -value =  $1.39 \times 10^{-15}$ ). Further, six of these genes (*DDX3*, *EIF1AX*, *EIF2S3*, *ZRSR2*, *PRKX* and *ZFX*) have previously been reported as overexpressed in female lymphoblastoid cell lines.<sup>12</sup> However, none of the 12 differentially expressed genes harbour Xa-specific methylation patterns identified in blood or buccal cells.

In contrast, we found multiple probes showing Xa-specific methylation pattern in both buccal and whole-blood tissues at the *BCOR* locus. However, these differences did not correlate with measurable differential gene expression in fractionated blood samples. Taken together, these results may reflect a role for Xa constitutive methylation in processes other than sex-specific gene expression. However, we cannot rule out that *BCOR* Xa methylation has a role in sexual dimorphic gene expression in specific cell types or below the detection level of our sample size. We note that *BCOR* is subject to recurrent somatic point mutations<sup>24,25</sup> or translocations<sup>28,29</sup> in a range of cancer sub-types.

To investigate sex-specific gene expression associated with our Xa probes, we performed a global analysis of gene expression of all genes nearby to the Xa probes. To do this, we first selected all genes located

within 2 kb of the Xa probes in both tissues and compared the expression profiles between males and females. We found 132 genes and miRNAs in this category and 107 of these genes had their expression profile available from the above data set. The *XIST* transcript again showed the strong sex-specific gene expression. In addition, the *MAGEA8* transcript was slightly over-expressed in females than males but only in monocytes (data not shown).

#### Xa-specific methylation is partially concordant with Turner syndrome Xa-sites

Sharp and co-workers<sup>14</sup> have previously examined XCI in peripheral blood comparing methylation levels of seven unrelated female patients carrying a single X chromosome (defined by the 45,XO karyotype characteristic of Turner syndrome) with three unrelated control female samples. The study reported methylation profiles predictive of XCI escape, identifying a total 313 candidate genes<sup>14</sup> with either low methylation ( $<0.65$ ) in females or low difference ( $<0.39$ ) between Turner syndrome cases and control females.

Comparison of our Xa constitutive genes (total 94 genes) with these candidate XCI escapee genes revealed an overlap, with 16 genes (17%). We note that Sharp *et al* used a higher density method to interrogate X-chromosome methylation, assaying more than one million probes compared with the HM450 array, which we used to assay 9558 probes. Interestingly, of the 10 genes we identified as upregulated in females (Supplementary Figure 4), only a single gene (*PRKX*) was not listed as a candidate XCI escapee. Thus, we found evidence that sex-specific gene expression is associated with Xa methylation. However, a large number of Xa constitutive sites were not associated with gene expression differences.

As the age of the individuals and tissue types were not matched between ours and the public data set, we cannot make a firm conclusion on the role of Xa-specific methylation on modulating gene expression. Nevertheless, given the complex relationship between gene expression and DNA methylation marks,<sup>30</sup> this result is not so surprising. It is possible that some of these methylation marks may act *in trans* or through more complex molecular pathways, especially as enhancers have been often shown to act this way.<sup>31</sup> Further investigation is required for more complete understanding of this relationship.

In the general absence of a link between Xa constitutive methylation marks and sex-specific gene expression, we sought to identify an

overlap between Xa methylation and sex-specific active histone modification patterns. An accumulation of active marks in females but not males would point to more accessible chromatin and a potential for expression on the inactive X. To do this, we used a publicly available active histone track (H3K4me3) from three male and three female monocyte samples, provided by the BLUEPRINT project in the UCSC Genome Browser.<sup>32</sup> Again, as expected, *XIST* showed a strong H3K4me3 signal at the promoter region, accompanied by strong expression in females only, as measured by RNA-seq (Supplementary Figure 5A). Another two regions, covering *LOC286467* and near *LOC642776*, showed overlap between Xa probes and H3K4me3, both with an absence of gene expression differences (Supplementary Figures 5B and C). Remarkably, *LOC286467* was previously identified as the only locus besides *XIST* with sex-specific RNA polymerase 2 binding,<sup>33</sup> and along with *XIST*, *LOC550643* was the only locus with sex-specific DNase I hypersensitivity patterns.<sup>34</sup> Unfortunately, the Infinium HM450 array did not cover the gene body of *LOC550643*, which showed H3K4me3 enrichment in females from the BLUEPRINT data. Our analysis therefore shows that Xa-specific methylation is associated with loss of active histone marks in males at genomic regions that have previously been implicated in sex-specific roles in X inactivation.<sup>34</sup> Therefore, we provide preliminary evidence for a role of DNA methylation on the active X in maintaining sex-specific accessible-chromatin patterns. However, this appears limited to only a few specific loci, and is not a widespread phenomenon.

## CONCLUSIONS

There are many phenotypic differences between sexes in mammals, many of which are attributable to the cascade of events associated with the presence of key sex determining genes on the Y chromosome. However, there are also many sex-specific effects that are poorly understood. Previous data suggested that a small proportion of genes on the X chromosome escape XCI by DNA methylation, a mechanism potentially responsible for controlling sex-specific dimorphisms.

Here we report 170 probes, linked to 94 unique genes that show Xa-specific hypermethylation. In this present study, we were not able to identify any functional implications these methylation marks are associated with. However, more complex and indirect regulation of gene functions from these marks could be possible, although we cannot rule out an expression-neutral effect. It may be possible that some of these marks are involved in the packaging of active X chromosome. Future studies are required to determine the extent of such methylation across a wider range of tissues in order to determine the functional consequences of such epigenetic marks on global gene transcription levels and content (eg, splicing variants) between sexes.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

This study is supported by the National Health and Medical Research Council (Project Grant 491246 to LWD; 437015 and 607358 to JMC and RS; Centre for Clinical Research Excellence 546519 to LWD), by the MCRI and the Victorian Government's Operational Infrastructure Support Program.

1 Wutz A: Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat Rev Genet* 2011; **12**: 542–553.

- 2 Payer B, Lee JT, Namekawa SH: X-inactivation and X-reactivation: epigenetic hallmarks of mammalian reproduction and pluripotent stem cells. *Hum Genet* 2011; **130**: 265–280.
- 3 Takagi N, Sasaki M: Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* 1975; **256**: 640–642.
- 4 Morey C, Avner P: The demoiselle of X-inactivation: 50 years old and as trendy and mesmerising as ever. *PLoS Genet* 2011; **7**: e1002212.
- 5 Brown CJ, Lafreniere RG, Powers VE et al: Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* 1991; **349**: 82–84.
- 6 Brown CJ, Ballabio A, Rupert JL et al: A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 1991; **349**: 38–44.
- 7 Brockdorff N: Chromosome silencing mechanisms in X-chromosome inactivation: unknown unknowns. *Development* 2011; **138**: 5057–5065.
- 8 Brockdorff N, Ashworth A, Kay GF et al: Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 1991; **351**: 329–331.
- 9 Jeon Y, Sarma K, Lee JT: New and Xisting regulatory mechanisms of X chromosome inactivation. *Curr Opin Genet Dev* 2012; **22**: 62–71.
- 10 Sado T, Hoki Y, Sasaki H: Tsix silences Xist through modification of chromatin structure. *Dev Cell* 2005; **9**: 159–165.
- 11 Carrel L, Willard HF: X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* 2005; **434**: 400–404.
- 12 Johnston CM, Lovell FL, Leongamornlert DA, Stranger BE, Dermitzakis ET, Ross MT: Large-scale population study of human cell lines indicates that dosage compensation is virtually complete. *PLoS Genet* 2008; **4**: e9.
- 13 Yasukochi Y, Maruyama O, Mahajan MC et al: X chromosome-wide analyses of genomic DNA methylation states and gene expression in male and female neutrophils. *Proc Natl Acad Sci USA* 2010; **107**: 3704–3709.
- 14 Sharp AJ, Stathaki E, Migliazza E et al: DNA methylation profiles of human active and inactive X chromosomes. *Genome Res* 2011; **21**: 1592–1600.
- 15 Cruickshank MN, Oshlack A, Theda C et al: Analysis of epigenetic changes in survivors of preterm birth 1 reveals the effect of gestational age and evidence for a long term legacy. *Genome Med* 2013; **5**: 96.
- 16 Martino D, Loke YJ, Gordon L et al: Longitudinal, genome-scale analysis of DNA methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance. *Genome Biol* 2013; **14**: R42.
- 17 Roberts G, Cheong J, Opie G et al: Growth of extremely preterm survivors from birth to 18 years of age compared with term controls. *Pediatrics* 2013; **131**: e439–e445.
- 18 Saffery R, Morley R, Carlin JB et al: Cohort profile: the peri/post-natal epigenetic twins study. *Int J Epidemiol* 2012; **41**: 55–61.
- 19 Ollikainen M, Smith KR, Joo EJ et al: DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. *Hum Mol Genet* 2010; **19**: 4176–4188.
- 20 Gentleman RC, Carey VJ, Bates DM et al: Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004; **5**: R80.
- 21 Maksimovic J, Gordon L, Oshlack A: SWAN: Subset quantile Within-Array Normalization for Illumina Infinium Human Methylation450 BeadChips. *Genome Biol* 2012; **13**: R44.
- 22 Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; **3**: Article3.
- 23 Davis S, Meltzer PS: GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* 2007; **23**: 1846–1847.
- 24 Grossmann V, Tiacci E, Holmes AB et al: Whole-exome sequencing identifies somatic mutations of BCOR in acute myeloid leukemia with normal karyotype. *Blood* 2011; **118**: 6153–6163.
- 25 Pugh TJ, Weeraratne SD, Archer TC et al: Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* 2012; **488**: 106–110.
- 26 Gianfrancesco F, Sanges R, Esposito T et al: Differential divergence of three human pseudoautosomal genes and their mouse homologs: implications for sex chromosome evolution. *Genome Res* 2001; **11**: 2095–2100.
- 27 Charchar FJ, Svartman M, El-Mogharbel N et al: Complex events in the evolution of the human pseudoautosomal region 2 (PAR2). *Genome Res* 2003; **13**: 281–286.
- 28 Pierron G, Tirode F, Lucchesi C et al: A new subtype of bone sarcoma defined by BCOR-CCNB3 gene fusion. *Nat Genet* 2012; **44**: 461–466.
- 29 Yamamoto Y, Tsuzuki S, Tsuzuki M, Handa K, Inaguma Y, Emi N: BCOR as a novel fusion partner of retinoic acid receptor alpha in a t(X;17)(p11;q12) variant of acute promyelocytic leukemia. *Blood* 2010; **116**: 4274–4283.
- 30 Ndlovu MN, Denis H, Fuks F: Exposing the DNA methylation iceberg. *Trends Biochem Sci* 2011; **36**: 381–387.
- 31 Jones PA: Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012; **13**: 484–492.
- 32 Adams D, Altucci L, Antonarakis SE et al: BLUEPRINT to decode the epigenetic signature written in blood. *Nat Biotechnol* 2012; **30**: 224–226.
- 33 Reddy TE, Gertz J, Pauli F et al: Effects of sequence variation on differential allelic transcription factor occupancy and gene expression. *Genome Res* 2012; **22**: 860–869.
- 34 Sheffield NC, Thurman RE, Song L et al: Patterns of regulatory activity across diverse human cell types predict tissue identity, transcription factor binding, and long-range interactions. *Genome Res* 2013; **23**: 777–788.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)