28. Medico, E. *et al.* The tyrosine kinase receptors Ron and Sea control "scattering" and morphogenesis of liver progenitor cells *in vitro*. *Mol. Biol. Cell* **7,** 495–504 (1996).
29. Li, C. & Wong, W. H. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl Acad. Sci. USA* **98,** 31–36 (2001).
30. Oshima, M. *et al.* Chemoprevention of intestinal polyposis in the Apcdelta716 mouse by rofecoxib, a specific cyclooxygenase-2 inhibitor. *Cancer Res.* **61,** 1733–1740 (2001).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank M. Risio, E. David and their staffs for pathology, M. Cilli for mouse surgery, R. Albano, A. Ferraro and R. Lo Noce for technical assistance, and M. Belluardo for computational analysis. We thank A.Cignetto for secretarial assistance. G.S. is the recipient of an AIRC fellowship. This research was supported by AIRC, CNR-MIUR, FIRB-MIUR, MIUR-PRIN and the Foundations CRT and 'Compagnia di San Paolo'.

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

Correspondence and requests for materials should be addressed to C.B. (carla.boccaccio@ircc.it).

X-inactivation profile reveals extensive variability in X-linked gene expression in females

Laura Carrel[1] **& Huntington F. Willard**[2]

[1]*Department of Biochemistry and Molecular Biology, Pennsylvania State University College of Medicine, Hershey, Pennsylvania 17033, USA*
[2]*Institute for Genome Sciences & Policy, Duke University, Durham, North Carolina 27708, USA*

In female mammals, most genes on one X chromosome are silenced as a result of X-chromosome inactivation[1,2]. However, some genes escape X-inactivation and are expressed from both the active and inactive X chromosome. Such genes are potential contributors to sexually dimorphic traits, to phenotypic variability among females heterozygous for X-linked conditions, and to clinical abnormalities in patients with abnormal X chromosomes[3]. Here, we present a comprehensive X-inactivation profile of the human X chromosome, representing an estimated 95% of assayable genes in fibroblast-based test systems[4,5]. In total, about 15% of X-linked genes escape inactivation to some degree, and the proportion of genes escaping inactivation differs dramatically between different regions of the X chromosome, reflecting the evolutionary history of the sex chromosomes. An additional 10% of X-linked genes show variable patterns of inactivation and are expressed to different extents from some inactive X chromosomes. This suggests a remarkable and previously unsuspected degree of expression heterogeneity among females.

Establishment of X-inactivation requires a key *cis*-acting master locus that includes the non-coding *XIST* gene[6,7]. How *XIST* RNA and other epigenetic modifications[8–10] are directed to sites along the inactive X chromosome (X_i), and how inactivation spreads in *cis* over the 155 Mb X chromosome still remain poorly understood. That some X-linked genes should escape inactivation was first predicted for pseudoautosomal genes on the X and Y chromosomes that show equivalent dosage, with two active copies in both males and females[11]. Because genes that escape inactivation lack at least some of the epigenetic alterations that characterize the rest of the chromosome[12], their identification and analysis are important for understanding the mechanics of the X-inactivation process.

We have previously developed an X-inactivation assay system to detect gene expression directly from the X_i in human female cells by determining relative expression levels of polymorphic alleles in heterozygous fibroblasts that demonstrate non-random inactivation[5]. For this study, we developed a quantitative assay based on fluorescent, single-nucleotide primer-extension (Q-SNaPshot) to distinguish between alleles of expressed single-nucleotide polymorphisms (SNPs) along the X chromosome (see Methods). For each heterozygous human fibroblast sample, complementary DNA was amplified, and the relative expression of each allele was determined after normalization to genomic DNA amplification products (Fig. 1). As all of the samples used show complete non-random inactivation[5], genes that show biallelic expression must escape inactivation, with the normalized allelic ratios indicating the level of X_i expression (relative to that on the active X chromosome, X_a), for example, *HDHD1A* and *RIBC1* in Fig. 1. In contrast, monoallelic expression indicates that a gene is subject to inactivation (for example, the *TMLHE* gene in Fig. 1). Ninety-four genes spanning the X chromosome were tested in a total of 40 human samples (Supplementary Table 1). Notably, only about 65% of the genes were subject to inactivation in all heterozygous samples (Fig. 2); a further 20% were inactivated in some, but not all samples, and 15% escaped inactivation in all samples. Surprisingly, among

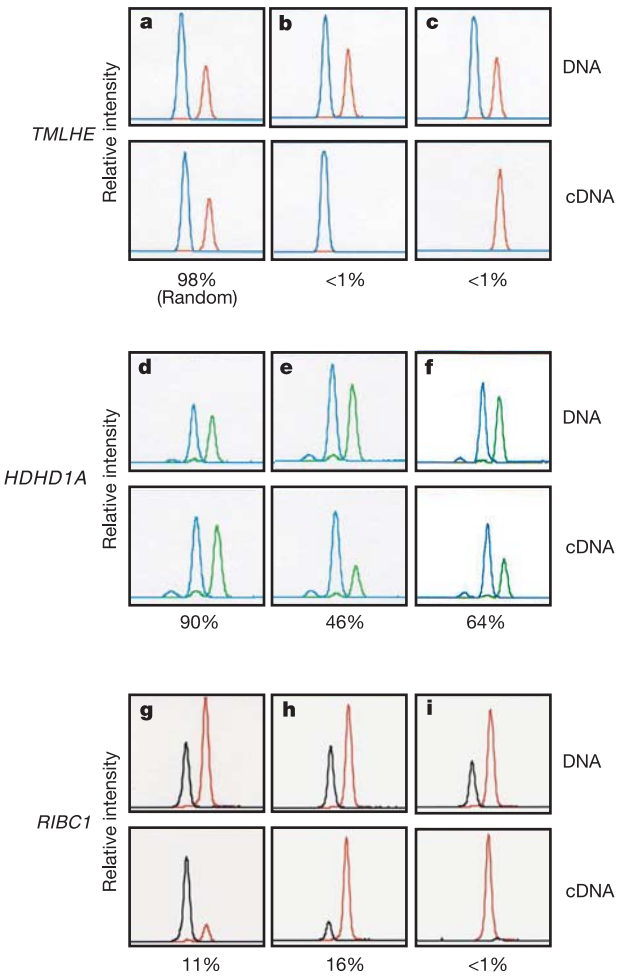

Figure 1 Q-SNaPshot assay of allelic expression for three X-linked genes. The assay measures expression by incorporating fluorescent dideoxy-nucleotides at polymorphic sites: G (blue), A (green), T (red), C (black). For each gene, DNA and cDNA amplification products are shown for three different cell lines (denoted **a**–**i**). For *TMLHE*, a randomly inactivated normal fibroblast line is included, whereas all other samples are non-randomly inactivated primary lines (that is, the same X chromosome is the X_i in all cells). Relative expression of the less intense allele for each sample (normalized with DNA to compensate for biased fluorescence output) is indicated as a percentage of the more intense allele, and by inference reflects the X_i expression level.

those genes that escape inactivation, most were not fully expressed from the X_i, demonstrating that escape from inactivation, even for pseudoautosomal genes, is generally partial and incomplete. For example, although the *FLJ39679* gene showed substantial X_i expression (>75% of Xa expression) in some cell lines, other cell lines showed X_i expression levels as low as ~25% of X_a expression levels (Fig. 2). These data indicate that females have considerable heterogeneity in levels of X-linked gene expression. At this point, however, we cannot establish whether this allelic variation is a specific feature of X-inactivation or reflects genome-wide variation.

To complement these data on a more comprehensive scale, we have established an X-inactivation profile in parallel with the complete sequence of the human X chromosome[13]. We determined the X-inactivation status of essentially all X-linked transcripts that are expressed in fibroblast-based assay systems[4,5] (see Methods and Supplementary Table 2), and conclude that the current survey of 624 transcripts represents ~95% of assayable genes on the X chromosome.

X-inactivation status was assessed by analysing gene expression from human X_a and X_i chromosomes in an extensive panel of rodent/human somatic cell hybrids[4]. This method does not require distinction between X_a and X_i transcripts, and is applicable to all X-linked genes expressed in these cells. Each transcript was assayed using polymerase chain reaction (PCR) with reverse transcription (RT–PCR) to determine whether expression is detected only from hybrid lines retaining X_a chromosomes (and is therefore a gene subject to X-inactivation) or from both X_a hybrid lines and X_i hybrid lines, indicating a gene that escapes X-inactivation. Genes were assayed in up to nine rodent/human somatic cell hybrids, each containing a different, normal human X_i chromosome (see Methods). Results for each gene in each X_i hybrid tested are depicted in Fig. 3a and are summarized for each transcript in Supplementary Table 3.

A total of 401 X-linked transcripts gave completely concordant results in all hybrids; that is, they were either expressed (74 transcripts) or silenced (327 transcripts) in all X_i hybrids tested. Considering only these transcripts (and excluding pseudoautosomal genes), the frequency of escape from inactivation for X-specific transcripts is minimally 16%, essentially identical to

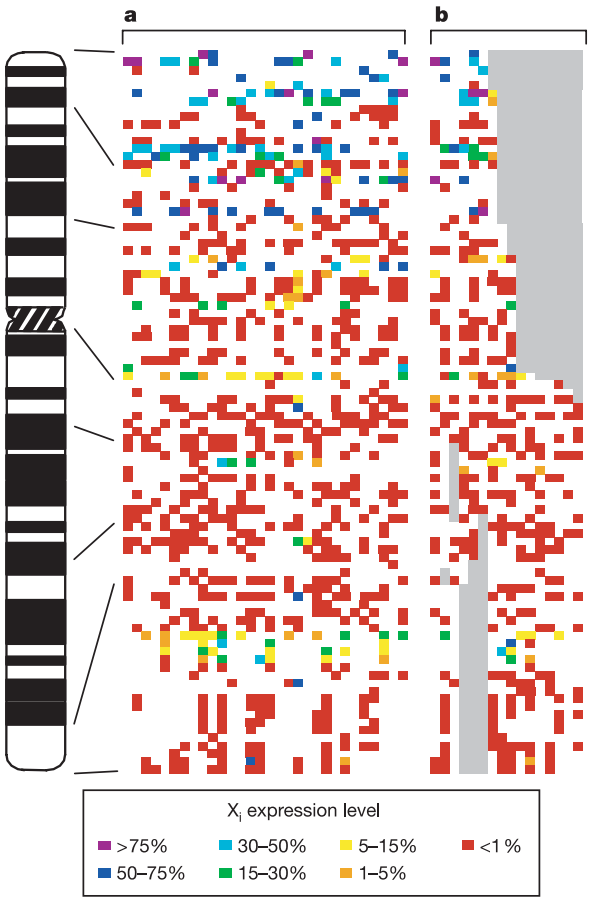

Figure 2 X-inactivation as a measure of allelic expression in non-randomly inactivated primary fibroblasts. **a**, **b**, Gene expression levels from normal X_i chromosomes (**a**) and from structurally abnormal X_i chromosomes (**b**). Each gene is linearly arrayed and approximate correspondence to chromosome location is indicated. The gene order is as in Supplementary Table 1. Each gene was assayed in all heterozygous individuals. Uninformative samples remain uncoloured. Colour-coding reflects relative X_i expression level as indicated. Grey boxes in **b** indicate absent portions of the X chromosome due to deletions or translocations.

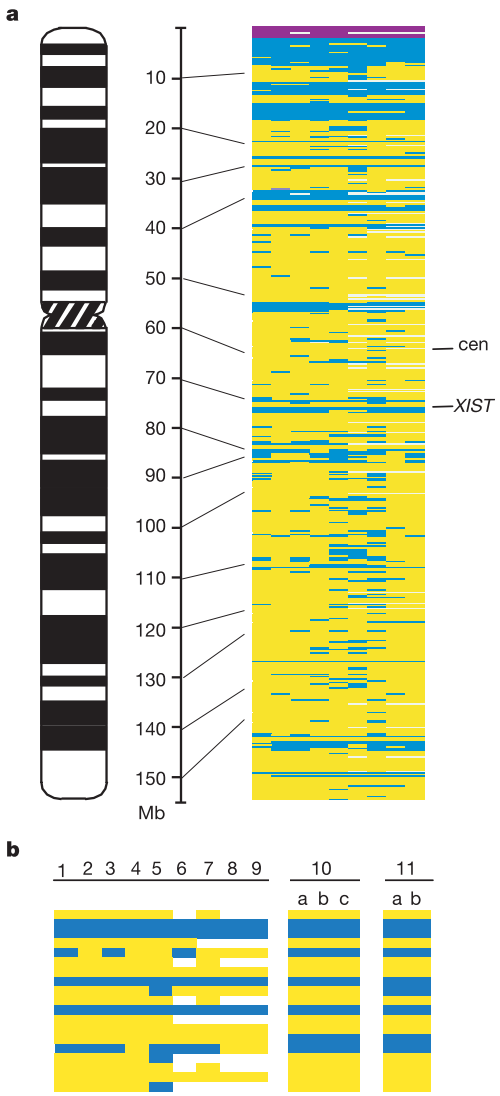

Figure 3 X-inactivation profile of the human X chromosome. **a**, 624 genes were tested in nine X_i hybrids. Each gene is linearly displayed. Blue denotes significant X_i gene expression, yellow shows silenced genes, pseudoautosomal genes are purple, and untested hybrids remain white. Positions of the centromere (cen) and *XIST* are indicated. **b**, To determine whether heterogeneity is largely a property of a specific chromosome, three independently derived hybrids from one X_i chromosome (denoted 10 a, b, c) and two hybrids carrying another X_i chromosome (11 a, b) were isolated, and results for 19 genes are shown adjacent to results for the original nine X_i chromosomes.

the estimate derived from studies in human cells. An additional 223 genes showed heterogeneity among different X_i hybrids and were expressed in some, but not all, of the X_i hybrids assayed (Fig. 3a), confirming and extending our observations in human cells (Fig. 2). Although this assay is inherently less quantitative then the SNP-based assay (see Methods), the majority of the genes escaping inactivation showed robust X_i expression; a minority were expressed from the X_i chromosome at relatively low levels (~10–15% that of the X_a). To assess the chromosomal basis for X_i heterogeneity, we analysed the same X chromosome isolated independently in multiple hybrids (Fig. 3b). Although the data set is limited, no variability was observed between different hybrids carrying the same X_i, supporting the idea that heterogeneity of X_i expression is largely an intrinsic property of the X chromosome tested, rather than an artefact of the hybrid model system. Of these heterogeneous genes, the majority (161 transcripts) gave concordant expression patterns in all but one or two hybrids, and for classification purposes[4], were considered to escape (30 genes) or be subject to (131 genes) X-inactivation. In total, of the 612 X-specific transcripts classified in this manner, 458 (75%) are subject to inactivation in most or all of the X chromosomes tested, and 94 (15%) escape inactivation; the remaining 10% are heterogeneous (Fig. 3a and Supplementary Table 3). Combined, these last two categories of genes indicate that at least 25% of X-linked genes escape inactivation in all or a significant subset of X chromosomes tested, a result fully consistent with the data obtained with human cells (Fig. 2).

Importantly, given the limited number (<50) of X chromosomes screened in the two assays, the results from the primary human cell assay were in agreement with the X_i hybrid expression patterns in most instances. For example, 17 of 20 genes that escaped inactivation in X_i hybrids were expressed from the X_i in at least some of the human samples tested, although only 12 of these genes were expressed in all heterozygous lines tested (Supplementary Table 4). This comparison suggests that the total proportion of X-linked genes showing heterogeneous patterns of X_i expression in the general population may be even greater than the estimates provided here.

The distribution of genes that escape inactivation is clearly not random along the chromosome[14]. Large blocks of transcripts expressed from the X_i, apparent in both assays (Figs 2, 3), indicate that these genes are clustered and map primarily to the distal portion of the X chromosome short arm (Xp), perhaps related to their distance from the *XIST* gene, as recently proposed for the mouse[15]. Mammalian X and Y chromosomes are proposed to have diverged from an identical pair of ancestral autosomes[16]. Key to their evolution, structural rearrangements successively suppressed recombination, allowing independent evolution of each chromosome and leading to the largely degraded Y chromosome, which contains very few genes[17]. Evolution of the X chromosome, on the other hand, has been heavily influenced by X-chromosome inactivation, and such stringent regulation requires that gene content remain highly conserved[16]. As a consequence, the long arm (Xq) of the human X chromosome and a portion of proximal Xp, retain a high degree of synteny even with marsupial and monotreme X chromosomes[18]. By using the boundary of this X-conserved region and calculating sequence divergence between X and Y homologues, the locations of these ancestral rearrangements have been inferred[19]. The X-specific portion of the chromosome can be partitioned into five evolutionary strata that show increasing levels of sequence divergence with increasing distance from the distal tip of Xp[13,19]. To explore an evolutionary influence on the X-inactivation profile, we compared X-inactivation patterns of genes mapping to each stratum (Fig. 4). Fulfilling the prediction that genes acquire the ability to be X-inactivated in response to the decay of Y-linked homologues[20], the proportion of genes that is expressed from the X_i in either the human samples (Fig. 4a) or in most hybrids (Fig. 4b) is clearly highest in the youngest stratum (stratum 5) and lowest in the oldest stratum (stratum 1).

Even excluding pseudoautosomal genes, the majority of X-linked genes with Y homology escape X-inactivation and are expressed in all X_i hybrids tested (22 out of 35), and this proportion is even higher when considering X-linked genes with functional Y orthologues. Notably, however, two such genes—*TSPYL2* and *RBMX*—were subject to inactivation in all hybrids; these genes belong to the previously defined ampliconic class of XY-homologous genes[17], suggesting that they may have less stringent dosage controls. We further examined the distribution and frequency of genes with Y homology, specifically among genes that were fully subject to or that fully escaped from inactivation (Supplementary Fig. 1). This comparison confirmed that younger strata contain a significantly higher proportion of genes with Y homology, but moreover, that the propensity of a gene with Y homology to escape inactivation diminishes in older strata, consistent with the idea that these genes lose the ability to escape inactivation over time[20].

The clustering of genes that escape inactivation, particularly evident in strata 4 and 5, supports the idea that control is at the level of chromosome domains, and is consistent with the premise that many genes escape inactivation because they lie within an epigenetic domain containing at least one X-linked gene with a Y homologue. However, this possibility cannot explain all genes that escape inactivation, because there are several other transcripts (particularly in stratum 1) that are expressed from the X_i even

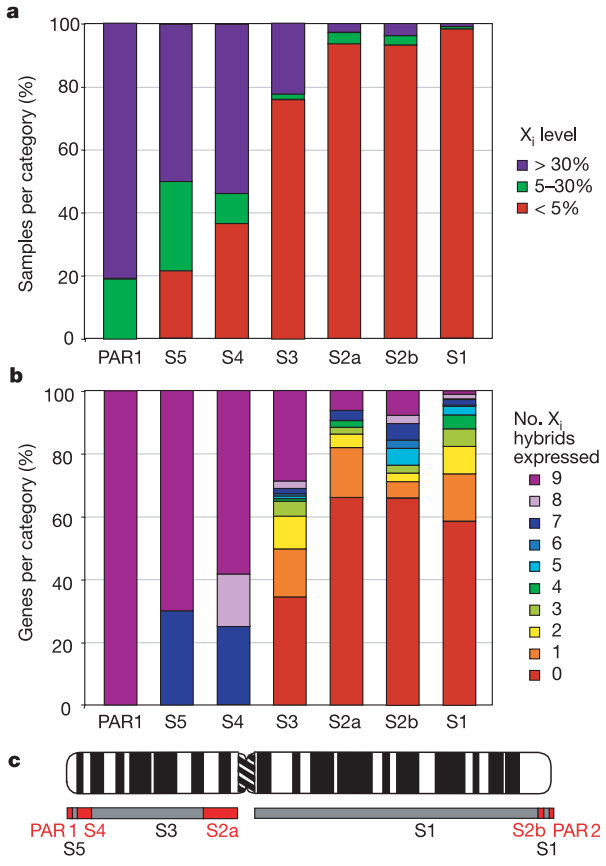

Figure 4 X_i expression data in primary fibroblasts and X_i hybrids correlate with location on X. Transcripts were subdivided according to X_i results and chromosomal region. **a**, The proportion of samples assayed in primary fibroblasts that show similar X_i results are indicated for the pseudoautosomal region PAR1, and for each X-specific evolutionary strata (S1–S5). **b**, For each chromosomal region, the percentage of genes showing similar X_i hybrid results (out of the nine X_i hybrid lines tested) are also shown. **c**, Location of each evolutionary strata on the X chromosome[13,19,21].

though they lack Y homology and map quite far from other expressed transcripts. One intriguing cluster of X_i-expressed genes maps within a gene-rich region of Xq28 that, on the basis of comparison with the chicken genomic sequence, has different evolutionary origins than the rest of Xq[21]. Using comparative mapping, this region is proposed to be included in stratum 2 (designated stratum 2b)[21], potentially explaining the number of genes in this region of Xq28 that are expressed from all or at least a high percentage of the X_i chromosomes tested (Figs 2, 3).

Models to explain how inactivation spreads across the X chromosome postulate that specific sequences along the chromosome function as *cis*-acting 'booster elements'[22]. Lyon has proposed that long interspersed repeat elements (L1) could serve such a function[23], in part because of their overall twofold enrichment on the X chromosome relative to autosomes[13,24]. L1 concentration is inversely correlated with the proportion of genes that escape inactivation[13] (Supplementary Fig. 2). Stratum 5 has the highest proportion of genes that escape inactivation, but also the lowest L1 and the highest *Alu* repeat concentrations. In fact, L1 levels in the pseudoautosomal region and in strata 4 and 5 are below the genome average[13]. Locally, however, differences in L1 concentration that correlate with X-inactivation patterns are not apparent (Supplementary Fig. 3). At the very least, these data indicate that the genomic and epigenomic determinants are likely to be more complex than repeat content alone.

Another local factor that could influence inactivation status is the presence of a CpG island, as DNA methylation is well known to be involved in the maintenance of X_i silencing[2,25]. In contrast to a previous report using a small number of genes[26], the whole chromosome profile shows that the distribution of island-associated genes does not differ between genes that escaped X-inactivation in all X_i hybrids and those that were completely silenced (Supplementary Fig. 4). However, the proportion of island-associated genes that are subject to or escape from inactivation does differ significantly from the proportion of genes containing CpG islands that demonstrate heterogeneous patterns of inactivation ($P < 0.001$ and $P < 0.025$ respectively). These data suggest that the presence of an island by itself does not determine whether a gene is inactivated; however, it seems that the absence of an island may contribute to the variability of X_i expression among different X chromosomes, as seen both in human cells and hybrids.

The large number of genes that escape inactivation, and their non-random distribution on the chromosome, has implications for counseling individuals with X chromosome abnormalities, estimated at 1 in 650 live births[3]. The evolutionary and genomic patterns of genes that escape inactivation suggest that aneuploidy for Xp is more severe than aneuploidy for Xq[4]. The profiles presented here provide a database of candidate genes to explain the clinical findings in such cases, although it is important to emphasize that the current profile has been obtained in fibroblast cells and may or may not reflect the situation in other tissues relevant to specific disease phenotypes.

Owing to different levels and different subsets of genes escaping X-inactivation, females will be even more variable with respect to X-linked gene expression than previously recognized. Because of these heterogeneous genes and the ~15% of genes that escape inactivation, the female genome differs from the male genome in at least four ways. First, the Y chromosome endows the male with at least several dozen genes that are absent in the female[17]. Second, the incomplete nature of X-inactivation means that at least 15% of X-linked genes are expressed at characteristically higher (but often variable) levels in females than in males. Third, a minimum of an additional 10% of genes show heterogeneous X-inactivation and thus differ in expression levels among females, whereas all males express a single copy of such genes. And fourth, the long-recognized random nature of X-inactivation indicates that females, but not males, are mosaics of two cell populations with respect to X-linked gene expression[1,3].

Notwithstanding the genomic and biological significance of these sex-specific differences, their general clinical and phenotypic implications remain unexplored. However, as many of the genes that escape from inactivation do not have Y-linked homologues, strict dosage compensation may not be necessary for all genes on the chromosome. Such characteristic genomic differences should be recognized as a factor for explaining sex-specific phenotypes both in complex disease as well as in normal, sexually dimorphic traits[27]. □

Methods

X-linked genes included in this study

Out of 931 X-linked genes annotated in NCBI build 34.3 (Supplementary Table 2), 267 were excluded from analysis because (1) they mapped to multiple locations, (2) they were computationally-predicted proteins that lacked significant expressed-sequence-tag (EST) support, or (3) they were transcripts with restricted expression patterns; for example, 10% of genes on the X chromosome are cancer-testis antigen genes that are not expressed in fibroblasts[13] (Supplementary Table 2 and Supplementary Note 1). Primers were designed for the remaining genes, and expression was initially tested in DNA and cDNA from human and mouse fibroblasts and mouse/human hybrid cell lines carrying X_a chromosomes. This preliminary screening identified 471 transcripts that were expressed in primary fibroblasts and could therefore be tested for X-inactivation. This frequency of expressed transcripts (71%) is consistent with microarray studies of fibroblasts[28]. An additional 153 transcripts (including full-length mRNAs and ESTs) that remain unassociated with currently annotated genes were also included in the profile. The transcripts analysed are listed in Supplementary Table 3.

Allelic expression in primary fibroblast cell lines

A panel of 40 primary fibroblast cell lines from females with non-random inactivation has been previously characterized[5]. These lines show complete skewing of inactivation, primarily as a result of secondary cell selection that maintains dosage despite structural rearrangements involving the X chromosome[3]. Confirmation that these lines were non-randomly inactivated was established by multiple assays[5]. Cell lines were maintained and DNA and cDNA isolated as described[5].

Several assays were used to determine relative allele frequencies[5]. A novel quantitative assay (Q-SNaPshot, a modification of the commercial SNaPshot assay (Applied Biosystems)) was established that uses primer extension to incorporate a single fluorescent dideoxy nucleotide at the polymorphic site and quantitates products on an ABI 3100 (L.C., unpublished). Samples were normalized by comparison to heterozygous DNA samples with a known 1:1 allele ratio (Fig. 1). Quantitative conditions were confirmed by correctly predicting allele ratios for known mixtures of homozygous samples (data not shown). All assays were performed in duplicate.

Expression of transcripts from active and inactive X hybrids

Somatic cell hybrids, cell culture conditions, and nucleic acid purification methods have previously been reported[4]. Initial studies established that many genes have very low levels of X_i expression (Supplementary Fig. 5), and we therefore adopted a protocol that would allow us to identify significant X_i expression levels in hybrids. Large-scale reverse transcription reactions were performed as described (but volumes were increased tenfold)[4]. RNA samples were treated with DNaseI before reverse transcription, to eliminate any contaminating genomic DNA. Each large-scale cDNA synthesis panel was normalized by amplification of the pseudoautosomal gene *MIC2*, which is equally expressed from X_a and X_i (Supplementary Fig. 5 and Supplementary Note 2).

Each gene was tested at two concentrations of cDNA, corresponding to ~50 ng and 250 ng of total RNA (Supplementary Fig. 6). Cycles were minimized to ensure that, in particular, the dilute cDNA samples were amplified in the exponential range. Although not strictly quantitative, this protocol allowed us to estimate the relative abundance of X_a and X_i transcripts. Faint amplification products that disappeared in the diluted samples were not considered positive. Samples that showed heterogeneous expression among different X_i chromosomes, or discordant patterns between cDNA dilutions, were repeated to ensure that scored positives were significantly expressed. Because of the semiquantitative nature of the RT–PCR assay used, the X_i expression levels in samples scored as positive are estimated to be at least at 10–15% of X_a levels (Supplementary Figs 5, 6 and Supplementary Note 2).

To assess the reproducibility of results within the same gene, ~15% of genes were assayed independently and blindly using at least two primer pairs. Virtually all (97%) of the X_i hybrid samples scored were in complete agreement; those that did not agree usually had faint bands that were called positive in one case but not another. For three genes, discordant results from more than one hybrid suggest that the assayed transcripts are incorrectly linked in the Unigene database and belong to distinct transcripts.

A small number of genes included in this profile have been tested by other laboratories (for example, see ref. 29). Most were re-analysed in our more extensive set of hybrids. The data here include and extend data previously reported from our laboratory[4].

Transcript analysis

Transcripts were mapped physically along the X chromosome using the May 2004 assembly from the UCSC genome browser (http://genome.ucsc.edu). CpG islands, as

annotated in the UCSC genome browser, were assigned to each gene if they lay within 2 kb upstream or downstream of the 5′ end of the RefSeq transcript or Unigene cluster. ESTs and most single exon transcripts were excluded from this analysis (Supplementary Table 3).

Evolutionary strata were assigned as described[13,19,21], and specific gene designations are indicated in Supplementary Table 3 and Supplementary Note 3. *Alu* and L1 repeat information was as annotated in the UCSC genome browser.

Received 24 January; accepted 17 February 2005; doi:10.1038/nature03479.

1. Lyon, M. F. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **190,** 372–373 (1961).
2. Plath, K., Mlynarczyk-Evans, S., Nusinow, D. A. & Panning, B. *Xist* RNA and the mechanism of X chromosome inactivation. *Annu. Rev. Genet.* **36,** 233–278 (2002).
3. Willard, H. F. in *The Metabolic and Molecular Bases of Inherited Disease* 8th edn (eds Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Childs, B. & Vogelstein, B.) 1191–1221 (McGraw-Hill, New York, 2000).
4. Carrel, L., Cottle, A. A., Goglin, K. C. & Willard, H. F. A first-generation X-inactivation profile of the human X chromosome. *Proc. Natl Acad. Sci. USA* **96,** 14440–14444 (1999).
5. Carrel, L. & Willard, H. F. Heterogeneous gene expression from the inactive X chromosome: an X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others. *Proc. Natl Acad. Sci. USA* **96,** 7364–7369 (1999).
6. Brown, C. J. *et al.* A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349,** 38–44 (1991).
7. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for *Xist* in X chromosome inactivation. *Nature* **379,** 131–137 (1996).
8. Rougeulle, C. *et al.* Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol. Cell. Biol.* **24,** 5475–5484 (2004).
9. Chadwick, B. P. & Willard, H. F. Barring gene expression after *XIST*: maintaining facultative heterochromatin on the inactive X. *Semin. Cell Dev. Biol.* **14,** 359–367 (2003).
10. Chadwick, B. P. & Willard, H. F. Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *Proc. Natl Acad. Sci. USA* **101,** 17450–17455 (2004).
11. Lyon, M. F. Sex chromatin and gene action in the mammalian X-chromosome. *Am. J. Hum. Genet.* **14,** 135–148 (1962).
12. Brown, C. J. & Greally, J. M. A stain upon the silence: genes escaping X inactivation. *Trends Genet.* **19,** 432–438 (2003).
13. Ross, M. T. *et al.* The DNA sequence of the human X chromosome. *Nature* doi:10.1038/nature03440 (this issue).
14. Miller, A. P. & Willard, H. F. Chromosomal basis of X chromosome inactivation: identification of a multigene domain in Xp11.21-p11.22 that escapes X inactivation. *Proc. Natl Acad. Sci. USA* **95,** 8709–8714 (1998).
15. Huynh, K. D. & Lee, J. T. Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* **426,** 857–862 (2003).
16. Ohno, S. *Sex Chromosomes and Sex-linked Genes* (Springer, Berlin, 1967).
17. Skaletsky, H. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423,** 825–837 (2003).
18. Graves, J. A. Mammals that break the rules: genetics of marsupials and monotremes. *Annu. Rev. Genet.* **30,** 233–260 (1996).
19. Lahn, B. T. & Page, D. C. Four evolutionary strata on the human X chromosome. *Science* **286,** 964–967 (1999).
20. Jegalian, K. & Page, D. C. A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. *Nature* **394,** 776–780 (1998).
21. Kohn, M., Kehrer-Sawatzki, H., Vogel, W., Graves, J. A. & Hameister, H. Wide genome comparisons reveal the origins of the human X chromosome. *Trends Genet.* **20,** 598–603 (2004).
22. Gartler, S. M. & Riggs, A. D. Mammalian X-chromosome inactivation. *Annu. Rev. Genet.* **17,** 155–190 (1983).
23. Lyon, M. F. X-chromosome inactivation: a repeat hypothesis. *Cytogenet. Cell Genet.* **80,** 133–137 (1998).
24. Bailey, J. A., Carrel, L., Chakravarti, A. & Eichler, E. E. Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proc. Natl Acad. Sci. USA* **97,** 6634–6639 (2000).
25. Mohandas, T., Sparkes, R. S. & Shapiro, L. J. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* **211,** 393–396 (1981).
26. Ke, X. & Collins, A. CpG islands in human X-inactivation. *Ann. Hum. Genet.* **67,** 242–249 (2003).
27. Committee on Understanding the Biology of Sex and Gender Differences. In *Exploring the Biological Contributions to Human Health–Does Sex Matter?* (eds Wizemann, T. M. & Pardue, M.) (National Acadamies Press, Washington, DC, 2001).
28. Iyer, V. R. *et al.* The transcriptional program in the response of human fibroblasts to serum. *Science* **283,** 83–87 (1999).
29. Esposito, T. *et al.* Escape from X inactivation of two new genes associated with DXS6974E and DXS7020E. *Genomics* **43,** 183–190 (1997).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank M. Ross for sharing information before publication and gratefully acknowledge technical assistance from G. Nickel, K. Trevarthen, J. Dunn, A. Cottle and M. Moon. This work was supported in part by a National Institutes of Health research grant to H.F.W. and L.C.

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

Correspondence and requests for materials should be addressed to L.C. (lcarrel@psu.edu).

The vacuolar Ca^{2+}-activated channel TPC1 regulates germination and stomatal movement

Edgar Peiter[1], Frans J. M. Maathuis[1], Lewis N. Mills[2], Heather Knight[3], Jérôme Pelloux[1], Alistair M. Hetherington[2] & Dale Sanders[1]

[1]Biology Department, Area 9, University of York, PO Box 373, York YO10 5YW, UK
[2]Biology Department, Lancaster Environment Centre, University of Lancaster, Lancaster LA1 4YQ, UK
[3]Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

Cytosolic free calcium ($[Ca^{2+}]_{cyt}$) is a ubiquitous signalling component in plant cells[1]. Numerous stimuli trigger sustained or transient elevations of $[Ca^{2+}]_{cyt}$ that evoke downstream stimulus-specific responses. Generation of $[Ca^{2+}]_{cyt}$ signals is effected through stimulus-induced opening of Ca^{2+}-permeable ion channels that catalyse a flux of Ca^{2+} into the cytosol from extracellular or intracellular stores. Many classes of Ca^{2+} current have been characterized electrophysiologically in plant membranes[2]. However, the identity of the ion channels that underlie these currents has until now remained obscure. Here we show that the *TPC1* ('two-pore channel 1') gene of *Arabidopsis thaliana* encodes a class of Ca^{2+}-dependent Ca^{2+}-release channel that is known from numerous electrophysiological studies as the slow vacuolar channel[3–5]. Slow vacuolar channels are ubiquitous in plant vacuoles, where they form the dominant conductance at micromolar $[Ca^{2+}]_{cyt}$. We show that a *tpc1* knockout mutant lacks functional slow vacuolar channel activity and is defective in both abscisic acid-induced repression of germination and in the response of stomata to extracellular calcium. These studies unequivocally demonstrate a critical role of intracellular Ca^{2+}-release channels in the physiological processes of plants.

In *A. thaliana* the *TPC1* gene (AGI code At4g03560) is the only representative of the TPC gene family[6] that is also present with one or two members in other dicotyledonous plants[7], monocotyledonous plants[8,9], gymnosperms (*Pinus pinaster*, GenBank accession number BX252040), mosses (*Physcomitrella patens*, GenBank accession number BQ039582), and animals[10]. The corresponding protein comprises two fused Shaker-like units, each with six transmembrane spans, a number of basic residues in the fourth transmembrane domain of each Shaker-like unit, a cytosolic linker region with two EF hands—that is, helix-loop-helix domains—suggestive of a role in Ca^{2+} binding (Fig. 1a), and several putative phosphorylation sites. Expression of plant TPC1 proteins in a yeast mutant lacking endogenous channel activity results in complementation of two phenotypic properties of the mutant—mating-induced cell death and low basal Ca^{2+} influx—suggesting that TPC1 might form a Ca^{2+}-permeable channel[6–9].

To determine the intracellular location of the TPC1 protein *in planta*, we constructed a TPC1–green fluorescent protein (GFP) fusion protein and introduced it into *Arabidopsis* mesophyll protoplasts. Figure 1b and c show images of protoplasts expressing *GFP* alone. In intact protoplasts, fluorescence appears as a clear band surrounding most of the cell, but is also present in a cytosol-rich region that surrounds the nucleus and chloroplasts. Fluorescence rapidly disperses after disruption of protoplasts expressing *GFP* alone, suggesting a cytosolic location for the GFP (Fig. 1c). Images of cells expressing the TPC1–GFP fusion protein are shown in Fig. 1d and e. The intact protoplast exhibits GFP fluorescence that is distinct from the plasma membrane and strongly suggestive of an association with intracellular membranes (Fig. 1d). Intact vacuoles