

Showcase on Research

Screening for Novel Mammalian Sex-Determining Genes using Expression Cloning and Microarray Approaches

Josephine Bowles, Monica Bullejos and Peter Koopman

Centre for Molecular and Cellular Biology, Institute for Molecular Biosciences,
University of Queensland, Brisbane, QLD 4072

Mammalian sex determination – the story so far

Disorders of sexual development are among the most common birth defects in humans. The infertility and sexual dysmorphism that accompanies these disorders is devastating for the individuals affected. At the root of these problems is a network of gene regulation that is responsible for making the decision to make testes or ovaries, for accurate programming of gonadal organogenesis, and for acquisition of the many secondary sexual characteristics that provide clear sexual identity and the ability to reproduce.

Maleness or femaleness in eutherian mammals is a matter of whether a testis or an ovary develops in the embryo (1). All other sex characteristics are thought to result from this decision, by means of factors produced by the gonads. In the mouse embryo, the gonadal primordia (genital ridges) first arise around 10.5 days *post coitum* (dpc), and sex determination is thought to occur around 11.5 dpc. The first visible sign of sex differentiation is the development and alignment of Sertoli cells into testis cords in males by about 12.5 dpc, which then leads to the differentiation of Leydig cells which synthesize steroid hormones, and to the formation of connective tissue and vasculature characteristic of the testes. In contrast, the ovaries do not overtly differentiate until about 16.5 dpc.

What do we know about the genes that control the sex of the embryo? It is now clear that testis formation is triggered by the Y-chromosomal gene *SRY* (2,3). *SRY* is expressed in the Sertoli cell lineage in male genital ridges during sex determination (10.5–12.5 dpc) (4). *SRY* protein contains a 79-amino acid motif (the HMG box), responsible for its ability to bind to known DNA sequences and to cause DNA to bend (5,6). Although *SRY* was discovered 10 years ago and has been studied intensively, its mode of action is still not clear and its target genes are yet to be identified.

A number of genes have been shown by

various means to play a role in mammalian gonadal development. These include the *SRY*-related gene *Sox9*, the orphan nuclear receptor gene *Dax1*, the Wilms tumour gene *WT1*, the zinc-finger gene *GATA4*, the polycomb-like gene *M33*, the DM domain gene *Dmrt1*, and genes encoding steroidogenic factor-1 (SF1) and anti-müllerian hormone (AMH). These genes are expressed in a sex-specific manner during gonadal development, and mostly encode transcription factors. In some cases, regulatory relationships have been elucidated (e.g. SF1 and SOX9 cooperate to regulate *Amh* expression), but for the most part the network of gene interaction remains unclear.

The field of sex determination is thus at a stage where a relatively small number of known genes is being studied intensely. Interactions between these genes and their products is being tested mostly by *in vitro* approaches. In many cases the *in vivo* significance of apparent interactions is not clear, and at least some may be artefacts resulting from attempts to reconstruct a complex puzzle using only a handful of available pieces.

Searching for the missing pieces of the sex-determination puzzle – our approach

It is clear that a large number of mammalian sex-determining genes remain to be identified. This theory is supported by detailed studies in the fruitfly *Drosophila* and the nematode worm *Caenorhabditis elegans*, by the high proportion of sex reversal and gonadal dysgenesis cases in humans for which no gene mutation has been identified, and by the cellular complexity of gonadal development. This regulatory network must consist of genes whose products are involved in cell-cell communication, signal transduction and transcriptional regulation, and it will be important to identify these genes in order to complete the molecular genetic picture of sexual development.

To this end, we have employed an expression-based gene discovery strategy

that utilizes suppression PCR and microarray technologies. Our approach is based on the simple premise that genes involved in sex determination will be expressed in the gonads during the sex-determining period, and that their expression profiles will differ between male and female gonads.

It has become clear that the success of microarray approaches rests on three main parameters: the quality of the probes, the quality of the arrays, and the quality of the screening process for evaluating positive hits. We have invested considerable effort in refining each of these parameters, to the point where a high proportion of primary hits appears to correspond to genes that genuinely differ in their expression between males and females.

In the following sections we describe our expression-screening strategy, which offers many advantages over alternative technologies, the most salient of which are:

- lack of bias regarding functional classes of output genes;
- sufficient sensitivity to allow detection of genes expressed at low abundance;
- bias towards detection of genes whose expression differs dramatically between the two samples being compared;
- outputs which, by definition, must be relevant to the biological system being studied;
- low rate of false positives; and
- rapidity and high throughput.

Production and evaluation of male- and female-enriched probes

Suppression PCR is a subtractive solution hybridization method that enriches for differentially expressed gene transcripts (7). The theoretical basis of the method is such that genes which are greatly up- or down-regulated (>5 fold) are more likely to be isolated than those showing only a moderate change in levels of expression (8). In addition, the suppression method results in normalization of the resulting cDNA pool - i.e. rare

Showcase on Research

Scening for Novel Mamalian Sex-Determining Genes (contin.)

species are relatively enriched compared to more abundant species with similar levels of differential expression. This method is rapidly gaining credibility through successful outcomes published in journals such as *Nature* (9) and *Cell* (10).

We used suppression PCR (PCR-Select, Clontech) to make male- and female-enriched cDNA samples, starting with pools of dissected and sexed fetal gonads. About 2 mg of poly(A)⁺ RNA from each tissue (in our case, fetal testes and fetal ovaries) is required to make the tester and driver cDNA in this technique. We used gonadal tissue from mouse embryos at 12.5-13.5 dpc, just after the testes have started to differentiate,

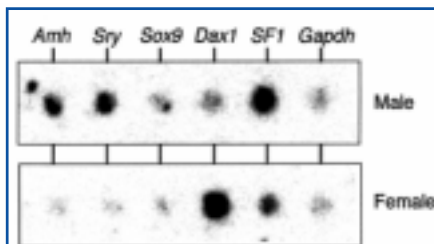


Fig. 1. Dot blot of known sex-determining genes, hybridized with male-enriched (upper panel) and female-enriched (lower panel) probes. *Sry*, *Sox9*, *Amh* and *SF1* were preferentially detected using the male-enriched probe, while *Dax1* was preferentially detected using the female-specific probe. The less clear-cut differences seen for *Sox9* and *SF1* may reflect the fact that some expression of these genes occurs in females at this time point, or that the subtraction method favours genes showing highly differential expression between the two tissues being compared.

so that the gonads could be sexed easily by morphological criteria.

To assess the efficiency of the subtractive approach, and evaluate the quality of our probes, we made duplicate dot-blot of genes known to be differentially expressed at this time-point. These blots were hybridized with radiolabelled male- and female-enriched probes. Results were in agreement with the known sexually dimorphic expression patterns of these genes (Fig. 1).

Further, we obtained strong signal for genes such as *Sry* which are expressed at very low levels in developing testes. These probes were also used to screen an Atlas Mouse cDNA array (Clontech). The male- and female-enriched probes gave differing

patterns of hybridization on the Clontech array. Results were dramatically clearer and cleaner than those previously obtained using conventional unsubtracted total cDNA probes. These data indicate that the probes were of high quality in terms of efficiency of subtraction, decreased complexity and sensitivity.

Identification of differentially expressed genes using conventional and microarray screening

The subtracted probes described above were cloned to produce subtracted male- and female-enriched cDNA libraries. We have used these libraries to generate candidate sex-determining genes in several different ways.

First, we have tested randomly-picked clones for sex-specific expression. We have found that a high proportion of these clones is expressed in gonads of one sex and not the other. These clones have included a high proportion of novel genes. Second, these libraries have been gridded onto nylon filters and probed with four different types of probe – male-enriched subtracted, female-enriched subtracted, male unsubtracted and female unsubtracted. Again, a large proportion of positive signals from these screen have proven to correspond to genes that show a genuine sex-specificity in expression (see below).

Greater efficiency can be achieved by using subtracted probes to screen micro-arrays. Clearly, microarray screening is revolutionizing gene discovery and functional genomics. We have deliberately avoided the use of whole genome-derived unigene sets in our microarray experiments. We reason that a large number of false positives is likely to arise from interrogating an array containing a vast number of genes that mostly have no relevance to the biological

system being investigated. Instead, we have tailored the micro-arrays to maximize the likelihood of discovering genes genuinely involved in gonadal development.

We are currently using custom-gridded microarrays of a mouse cDNA library prepared from mixed-sex 10.5-12.5 dpc urogenital ridge tissue, made by our collaborators, Dr Andy Greenfield (MRC Harwell, UK) and Dr Sean Grimmond (now at QIMR, Qld). This library has been normalized, gridded and sequenced by the I.M.A.G.E. Consortium, and is thus an excellent resource for our studies. We have fluorescently labelled our male-enriched and female-enriched probes and hybridized them, simultaneously, to these microarrays. Because of the normalization and enrichment inherent to our probes, a large number of strongly 'male-specific' and 'female-specific' clones have been identified (Fig 2).

Verification of a role in sex determination or sexual differentiation

Random selection of clones from a subtracted cDNA library, macroarray and microarray screening are methods that generate a large number of primary candidates for further evaluation. A high-

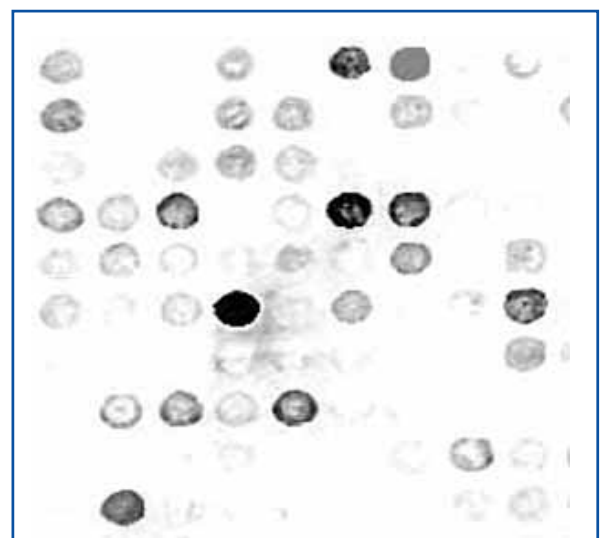


Fig. 2. Fluorescently-labelled subtracted probes used to screen a custom-made mouse urogenital ridge cDNA array. Male-enriched and female-enriched probes were hybridized simultaneously. Strong red signals correspond to male-specific candidates and strong green signals to female-specific candidates (not apparent on this black and white reproduction).

throughput screening method is required to sort the wheat from the chaff.

Showcase on Research

Scening for Novel Sex-Determining Genes (contin.)

sected male and female genital ridges at 11.5 and 12.5 dpc (Fig. 3).

We are aided in this endeavour by the availability of a robotic *in situ* hybridization machine capable of handling up to 96

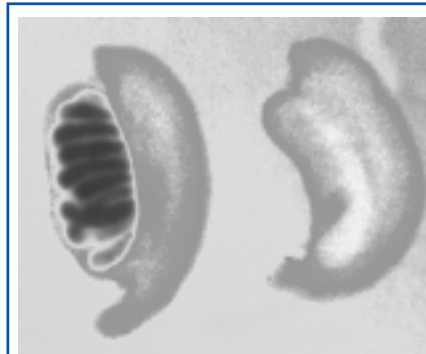


Fig. 3. Whole mount *in situ* analysis of the expression of one candidate sex determining gene, novel gene E5 in 12.5 dpc male and female urogenital ridges. The gene is highly expressed in the testicular cords of the male gonad (left), but is expressed at only very low levels in the female gonad (right).

samples simultaneously. To date, a surprisingly large proportion of primary candidates have shown sex-specific expression *in situ*.

For candidates that show sex-specific expression, further studies are then used to establish the timing of onset and the duration of expression. We are most interested in genes expressed by Sertoli cells, the cell type in which the sex-determining switch gene *Sry* functions. In the case of novel genes, we are also assessing what tissues other than developing gonads express each gene. This information will be an important factor in determining priorities and strategies for future knockout studies.

Of course, identifying genes expressed in a sex-specific manner by subtraction and microarray strategies is only half the fun. The real challenge is to tackle the biochemistry, molecular genetics and biology of these genes and their products. Novel genes identified as being sex-specifically expressed in the critical period of development will need to be fully cloned and sequenced in order to gain information about the likely biochemical class and hence role of the encoded proteins.

Together with data relating to the sex, stage and cell-type of gene expression, this information may suggest which other genes or gene products are likely to interact with each candidate. In the case of

growth factors, we intend to directly assay their role in gonadal development by applying these factors to cultures of mouse genital ridges. The effect on gonadal morphology will be assessed relative to controls of the same sex. Ultimately, the roles of individual genes can be assessed *in vivo* using gene targeting. In addition to examining the biological effects of gene inactivation on gonadal development, we will be able to repeat the cycle and exploit microarray technology to assess the molecular consequences of the knockout. In this way we will build up a picture of the gene regulatory cascade that programs gonadal development.

In summary, the type of approach we are applying to discovering genes involved in sexual development should be applicable to a wide range of biological problems, and offers the possibility of extending the already considerable power offered by microarray screening.

References

1. Jost, A. (1947) *Archs Anat. Microsc. Morph. Exp.* **36**, 271-315
2. Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A-M., Lovell-Badge, R., and Goodfellow, P.N. (1990) *Nature* **346**, 240-244
3. Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990) *Nature* **346**, 245-250
4. Koopman, P., Münsterberg, A., Capel, B., Vivian, N., and Lovell-Badge, R. (1990) *Nature* **348**, 450-452
5. Ferrari, S., Harley, V.R., Pontiggia, A., Goodfellow, P.N., and Lovell-Badge, R. (1992) *EMBO J.* **11**, 4497-4506
6. Harley, V.R., Jackson, D.I., Hextall, P.J., Hawkins, J.R., Berkovitz, G.D., Sockanathan, S., Lovell-Badge, R., and Goodfellow, P.N. (1992) *Science* **255**, 453-456
7. Diatchenko, L., Lau, Y.F., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E.D., and Siebert, P.D. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 6025-6030
8. Gurskaya, N.G., Diatchenko, L., Chenchik, A., Siebert, P.D., Khaspekov, G.L., Lukyanov, K.A., Vagner, L.L., Ermolaeva, O.D., Lukyanov, S.A., and Sverdlov, E.D. (1996) *Anal. Biochem.* **240**, 90-97
9. Yokomizo, T., Izumi, T., Chang, K., Takuwa, Y., and Shimizu, T. (1997) *Nature* **387**, 620-624
10. Hudson, C., Clements, D., Friday, R.V., Stott, D., and Woodland, H.R. (1997) *Cell* **91**, 397-405

