European Society of Human Reproduction and Embryology

COURSE 3

SIG Andrology course:

“The use of functional genomics in male fertility”

30 June 2002

Vienna - Austria

Contents

Course Program

Course Co-ordinators: Ch. De Jonge (USA) and Ch. Barratt (UK)

09.00-09.30	Introduction: The use of functional genomics in male infertility - ***Ch. Barratt (UK)***
09.30-09.45	Discussion
09.45-10.15	Bioinformatics : gene-protein-structure-function - ***T. Attwood (UK)***
10.15-10.30	Discussion
10.30-11.00	Coffee break
11.00-11.30	Gene expression in spermatogenesis using micro array technology - ***P. Turek (USA)***
11.30-11.45	Discussion
11.45-12.15	Sequencing they Y - what it means - ***R. Reijo (USA)***
12.15-12.30	Discussion
12.30-13.30	Lunch
13.30-14.00	KO mice - from mouse to man what can we learn - ***M. Eddy (USA)***
14.00-14.15	Discussion
14.15-14.45	Genes causing defects in spermatogenesis - ***D. DeKretser (AUS)***
14.45-15.00	Discussion
15.00-15.30	Coffee break
15.30-16.00	What can a structure tell us about function – acrosin - ***R. Jones (UK)***
16.00-16.15	Discussion
16.15-16.45	Application of the technology to patient management - ***J. Kremer (NL)***
16.45-17.00	Discussion
17.00-18.00	Special Interest Group Andrology business meeting

Functional Genomics in Male infertility - Introduction

1 Christopher J De Jonge and 2, 3 Christopher LR Barratt

1Reproductive Medicine Center
Department of Obstetrics & Gynecology
University of Minnesota
Minneapolis, MN 55454
Tel: 001 612 627 4807
Fax: 001 612 627 4334
Email: dejone@tc.umn.edu

2Assisted Conception Unit
Birmingham Women's Hospital
3 Reproductive Biology & Genetics Research Group
Department of Medicine
Medical School
University of Birmingham, UK
Tel & Fax: 0044 121 627 2797
Email: c.l.barratt@bham.ac.uk

Adapted from Barratt CLR et al (2002) Hum. Fert 5, 3-5; and De Jonge CJ & Barratt CLR (In Press) The Future of Reproductive Cellular Engineering in Male Infertility. Urology Clinics of North America.

Learning objectives :
1. Obtain a basic understanding of the range of functional genomic techniques relevant to male infertility.
2. Obtain a basic understanding of the advantages/disadvantages of these techniques and know several examples where these methods have been used to improve diagnosis.
3. Highlight several possible short term developments in male infertility that may rely heavily on using functional genomics.

INTRODUCTION

Currently it is almost impossible to open a scientific journal without an article describing the use of functional genomics. Dedicated journals are available with new ones being added. The aim of this article is to describe three of the main component of functional genomics using examples from the literature. Currently, there are relatively few publications using, for example, global gene expression profiles for germ cells so we have often used examples from other systems to provide a flavor of the potential uses of this technology.
Functional genomics can be split into three interrelated disciplines:

1. Bioinformatics

A variety of datasets are already available for interrogation including genomic sequences and expressed sequence tags (EST's). The elucidation of the human genome sequence (approximately 30,000-40,000 genes) provides an excellent focal point from which we can start to determine which genes may be

important for male fertility (ref). In fact, many recent discoveries have been made as a result of examining the currently available databases - for example - the fourth human zona pellucida gene. The availability of genomic sequences from close relatives to man (e.g. chimpanzee), to other mammals (mouse), plus numerous eukaryotic species, enables extensive comparison of, for example, coding sequences, and also the identification of common regulatory elements (see Fujiyama et al, 2002 for comparison between Chimpanzee and man). Such comparisons can provide a wealth of information - for example - in prokaryotes, comparative genomics has been used extensively to determine why some infections, e.g Listeria, prove to be fatal (Glaser et al, 2001).

2. Microarrays.

A major diagnostic development is the ability to analyze global gene expression profiles using microarrays (gene chips) (review - Young 2000). Microarrays can profile thousands of genes simultaneously and, by comparison with controls, can illuminate a series of defects. For example, microarrays with over 10,000 cDNA's were used to distinguish between neoplastic and normal prostatic tissue - allowing the first molecular profiling of prostate cancer (Dhanasekaran et al, 2001). Preliminary microarray analysis of Drosophila have been prepared that show a highly specialized expression profile with some genes being highly expressed in the testis yet minimally expressed elsewhere (e.g. head), presumably reflecting the high level of complexity required to remodel an undifferentiated stem cell into a mature spermatozoon (Andrews et al, 2000). Interestingly, a significant number of the predicted genes have no known function.

A 950 and 960 gene DNA array has been developed for the mouse and human testis respectively (Rockett et al, 2001). The availability of these arrays will provide an excellent starting point to investigate general and testis specific genes involved in spermatogenesis, as well as provide the starting point for the analysis of pathological samples, e.g. men with non-obstructive azoospermia.

Alternative technology for analyzing gene expression is available, e.g. SAGE (serial analysis of gene expression) which allows an unbiased sampling of the transcriptome. This is an iterative process in that SAGE data can be fed back into the design of microarrays. Using SAGE, the first expression profile of the human oocyte, detailing several thousand possible genes and gene families has been reported (see Nielson et al, 2000) and this may provide a complementary analysis to examining the specific genes expressed during spermatogenesis. Currently, this field is in its infancy but soon it is likely that gene expression profiling will be used to provide insights into the pathology of men with defective gametes and spermatogenesis.

3. Proteomics.

Another significant technological development has been in the area of Proteomics - the science of obtaining a comprehensive and systematic quantification and identification of proteins expressed in cells and tissues (see Griffin and Aebersold 2001). Considerable challenges still remain e.g. the under representation of low abundance proteins but these are likely to be overcome in the near future. John Herr and colleagues are already beginning the proteomic profile of human spermatozoa and, to date, have identified eight novel membrane proteins as potential contraceptive targets (see Shetty et al, 2001). There are likely to be many others but the technological limitations are currently limiting their identification. In addition, proteomics has been used to examine the proteome of spermatogonial germ cells. With developments in automation, it is likely that proteomic profiles of germ cell populations and mature sperm will be available in the very near future. One achieved, we can then have a clear insight into comparisons between pathological samples and cross-referencing with gene expression studies to allow identification of defective genes families/proteins.

Ultimately, of course, it may be possible to do high throughput screening of protein interactions using chip technology and astonishing progress is being made in this area suggesting that this will be possible in the near future. Preliminary results show that cellular pathways can be built and tested using microarray, proteomic and bioinformatic approaches (see Gavin et al., 2002).

Using proteomic profiling we may, in the future, be able to identify the cause of spermatogenesis dysfunction and provide treatment avenues for a man just by examining a serum sample. Proteomic patterns in the serum of ovarian cancer patients have been categorized and data is obtain within 30 minutes of sample collection (see Petricoin et al., 2002). This advance for male infertility would be breathtaking!

IN SUMMARY

It is critical that our understanding of the pathophysiology of defective spermatozoa and germ cell function be advanced so that appropriate rational treatment can be applied. Dramatic advances in ART have been made but the ultimate goal must be to develop such treatment so that the couple can conceive naturally. Functional genomics has allowed many advances in other disciplines. If used wisely and effectively such techniques are likely to provide quantum leaps in our understanding of the causes and thus potential treatment options for male infertility.

REFERENCES

Andrews J, Bouffard GG, Cheadle C, et al: Gene discovery using computational and microarray analysis of transcription in the Dropsophila melanogaster testis. Genome Res 10: 2030-2043, 2000

Dhanasekaran SM, Barrette TR, Ghosh D, et al: Delineation of prognostic biomarkers in prostate cancer. Nature 412(6849): 822-826 2001

Fujiyama A, Watanabe H, Toyoda A, et al: Construction and analysis of a human-chimpanzee comparative clone map. Science 295(5552): 131-134, 2002

Gavin AC, Bosche M, Krause R et al., Functional organization of the yeast proteome by systematic analysis of protein. Natuure 415, 141-147

Glaser P, Frangeul L, Buchrieser C, et al: Comparative genomics of Listeria species. Science 294: 849-852, 2001

Griffin TJ, Aebersold R: Advances in proteome analysis by mass spectrometry. Accepted J Biol Chem, 3 October 2001, published on WWW, 2001

Nielson L, Andalibi A, Kang D, et al: Molecular phenotype of the human oocyte by PCR-SAGE. Genomics 63: 13-24, 2000

Petricoin E.F., Ardekani, AM, Hitt, B A. Use of proteomic patterns in serum to identify ovarian cancer. The Lancet 359, 572-577

Rockett JC, Luft JC, Garges JB, et al: Development of a 950-gene DNA array for examining gene expression patterns in mouse testis. Genome Biol 2(4): 0014.1-0014.9, 2001

Shetty J, Diekman AB, Jayes FC, et al: Differential extraction and enrichment of human sperm surface proteins in a proteome:identification of immunocontraceptive candidates. Electrophoresis 22: 3053-3066, 2001

Young RA: Biomedical discovery with DNA arrays. Cell 102: 9-15, 2000

Bioinformatics: Gene-Protein-Structure-Function

Professor Teresa K. Attwood, PhD
School of Biological Sciences and Department of Computer Science
The University of Manchester
Oxford Road
Manchester M13 9PT

OVERVIEW

Despite considerable recent progress in computational biology, predicting genes in uncharacterised genomic DNA is still not a solved problem. De novo prediction methods (searching for splice-site consensus motifs, biased codon usage, and so on) have been only partially successful, and investigators have found that the surest way of predicting a gene is by alignment with a homologous protein sequence. This presentation will therefore focus on a discussion of alignment-based methods for identifying homologous relationships between proteins. Specifically, we will examine:

* in silico function prediction: how much functional information can we realistically ascertain from protein structures and sequences - a reality check;
* what analytical tools are available for sequence analysis (an overview of motif- and domain-based methods);
* what databases are used for family assignment (PROSITE, PRINTS, etc.);
* how to build a sequence analysis protocol, and why 'expert systems' should be handled with care; and
* what other integrated approaches are available.

IN SILICO FUNCTION PREDICTION

The importance of deducing the 3D structures of proteins cannot be disputed. One compelling argument is that structure tells us function. Sadly, however, this is not true. We can think of a protein fold as providing a scaffold, which can be decorated with different sequences to confer different functions. Knowing the fold and the function allows us to rationalise how the structure might effect its function at the molecular level. Looking at a structure in isolation, however, with no other background knowledge, it is almost impossible to say anything meaningful about its function - only if we've seen a structure before are we likely to be able to do this. Take, for example, the structure shown in Fig.1: this was hailed as a test-case for structural genomics, in a paper entitled "the structure-based assignment of the biochemical function of hypothetical protein MJ0577." Yet, on reading the small print, we learn that while the structure co-crystallised with ATP, its biochemical function remains unknown!

Fig.1. What is the function of this structure?

Interestingly, the same is true for sequences in isolation. We can take any sequence (e.g., Fig.2) and ask, what is its function? However, unless we've seen it before, it is unlikely that we will be able to deduce its function directly from its sequence.

Fig.2. What is the function of this sequence?

We begin to make progress if we ask, instead, what is the function of this motif? By aligning sequences into families, patterns of conservation emerge that begin to offer structural and functional clues. Thus we may look at the motif in Fig.3 and deduce from the conserved pattern of acidic residues that it might be involved in calcium binding.

Fig.3. What is the function of this motif?

The heart of sequence analysis, then, depends on building multiple alignments and searching for patterns of conservation that might offer structural or functional clues. If 3D coordinates are available, such patterns can be related directly to known structural features. However, there is a dearth of known structures by comparison with the abundance of available sequence data. Hence a variety of sequence analysis techniques have been devised to make structural and functional inferences directly from conserved regions observed in sequence alignments.

The challenge for family analysis

When attempting to deduce the biological functions of aligned protein families, one of the biggest hurdles is that computational tools know little or nothing of the underlying biology - sequences are just strings of symbols, without biological context or relevance.

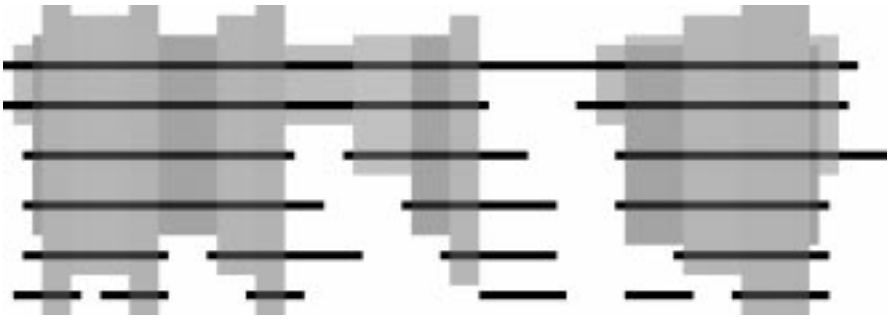


Fig.4. A discrete family, or several functionally distinct families in a superfamily?

Thus, a clustering algorithm might group together a range of increasingly divergent sequences, without offering any insight into the biological relationships between them. It is then for the user to determine whether, for example, the group of aligned sequences represents a discrete family with a single function, or whether it represents a superfamily that contains many related but functionally distinct families (see Fig.4). These different scenarios must be distinguished if function prediction is ever to be achieved reliably in silico.

Three main approaches have been devised for family, and hence functional, analysis of protein sequences (Fig.5). The first adopts the principle that the single most conserved region within an alignment is sufficient to characterise the family, and hence to infer the function of a query sequence. This approach is embodied in the PROSITE database, which uses regular expressions to encode patterns of conservation. The second school of thought takes the view that alignments tend to be characterised, not by one, but by several conserved regions, which together provide a 'fingerprint' or signature of family membership - this is the basis of the PRINTS database. And the final approach goes a step further by including not only the most conserved regions, but also the gapped regions between them. It is beyond the scope of this presentation to describe all of these methods in detail (for further information, see Readings). However, for the purposes of illustration, we will take a closer look at two of them.

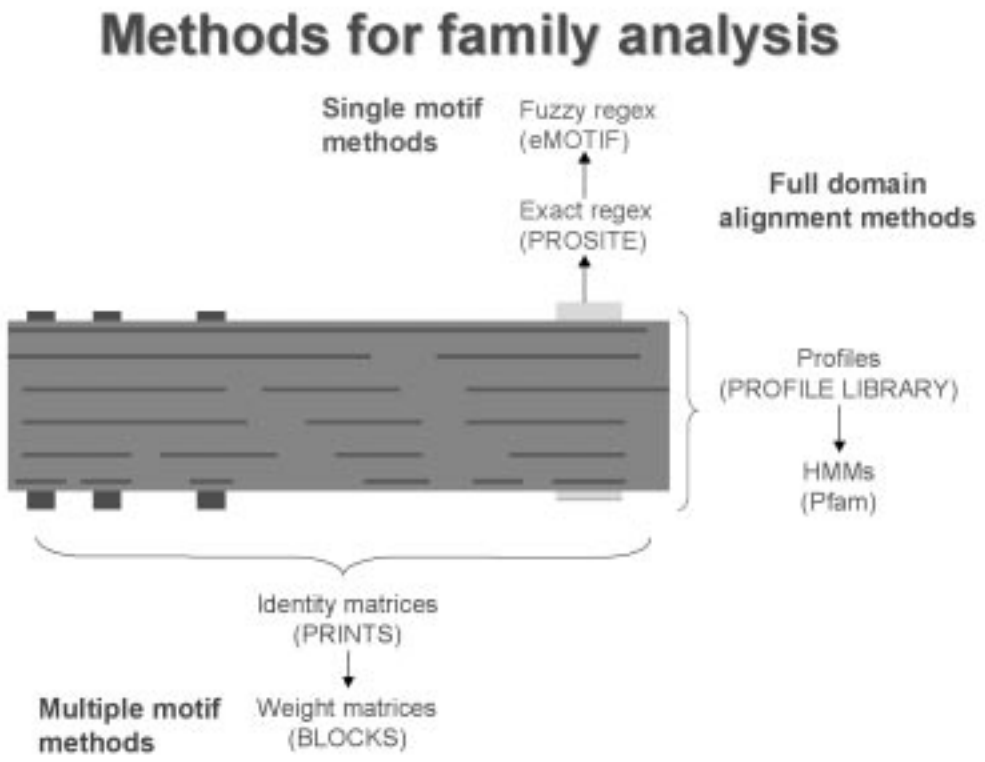


Fig.5. Comparison of single motif, multiple motif and domain-based analysis methods.

The PROSITE database

The first database that attempted to characterise protein families and functional sites was PROSITE, which encodes well-conserved motifs in the form of regular expression patterns. An excerpt from a PROSITE entry is shown in Fig.6:

ID G_PROTEIN_RECEPTOR; PATTERN
AC PS00237;
DE G-protein coupled receptor signature
PA [GSTALIVMYWC]-[GSTANCPDE]-{EDPKRH}-X(2)-[LIVMNQGA]-
PA X(2)-[LIVMFT]-[GSTANC]-[LIVMFYWSTAC]-[DENH]-R
NR /TOTAL=1121(1121); /POS=1057(1057); /FALSE_POS=64(64);
NR /FALSE_NEG=112; /PARTIAL=48; UNKNOWN=0(0)

Fig.6. Excerpt from PROSITE, showing a regular expression and its match statistics.

Within the pattern (PA lines), residues within square brackets are allowed at that position in the motif; those between curly brackets are disallowed at that position; x means any residue is allowed; and a residue on its own denotes a totally conserved position. The statistics of how well the pattern has performed are shown on the NR lines. Here, we see that, of 1121 matches, 64 were incorrectly matched and 112 were missed completely. The error rate for this entry is thus 20%, but since the data assume the curators haven't missed anything, the actual rate is probably higher. Thus, a match to a pattern isn't necessarily true, and a mismatch isn't necessarily false! False-negatives are a fundamental limitation to this type of pattern matching - in other words, if you don't know what you're looking for, you'll never know you missed it!

What is PRINTS?
To address these diagnostic limitations, protein fingerprinting was devised and gave rise to the PRINTS database. Fingerprints are groups of conserved motifs used for iterative database searching. The iterative process refines the fingerprint, and further potency is gained from the mutual context of motif neighbours - this means that results are biologically more meaningful than those from single motifs. In addition, the approach is hierarchical, allowing fine-grained diagnoses of family membership (see Fig.7).

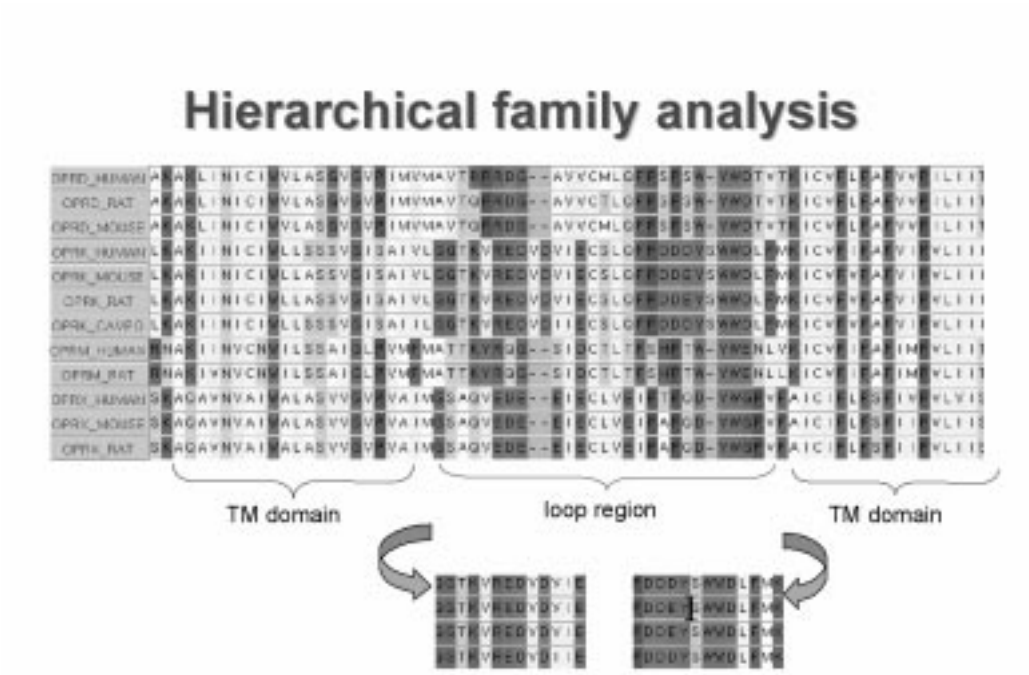


Fig.7. Illustration of protein fingerprinting showing excised motifs from a subfamily.

PRINTS and PROSITE are just two of several related protein family classification resources. In any search strategy, it is important to search a range of databases of this type, in addition to performing standard sequence database searches, because tools such as FastA and BLAST are not infallible. The hope, then, is that distant relationships missed, for example, by BLAST may be captured by one or more of the family or functional site distillations within the family classification resources.

Overview of resources
The most popular family classification resources are listed below: none is “best”, as each has its optimum area of application:

- PROSITE (1121 entries) - uses single motifs (regular expressions) and works best with small highly conserved sites;
- PRINTS (1700 entries) - uses multiple motifs (fingerprints) and works best for families and subfamilies;
- Profile library (300 entries) - uses weight matrices (profiles) and is good with divergent domains and superfamilies;
- Pfam (3621 entries) - uses hidden Markov models (HMMs) and is good with divergent domains and superfamilies;
- Blocks+ (2608 entries) - uses multiple motifs (blocks) derived from InterPro and PRINTS; and
- eMOTIF - uses fuzzy regular expressions derived from PRINTS and Blocks.

Designing a search protocol
Given a newly-determined sequence, we want to know, what is my protein? To what family does it belong? What is its function? And how can we explain its function in structural terms? Given the variety of analysis tools and databases available, rather than relying on just one, it is important to devise a search protocol that involves searches of both the sequence and family databases, and estimating the significance of matches by comparing results and finding a consensus. This does not simply mean running BLAST and PROSITE searches, which is easy to do on the Web, or FastA and motif/profile searches, such as are available in commercial packages like GCG. Nevertheless, this is still what most people do, including so-called ‘expert systems’ for genome analysis.

Expert systems for functional analysis
To facilitate the transition from genome data to biological knowledge, several expert systems have been devised. These include: GeneQuiz, for automatic protein function annotation; MAGPIE, for automatic genome analysis; and PEDANT, for automatic protein analysis.
If we look at how these systems describe themselves, we see that GeneQuiz is an expert system for derivation of functional information; MAGPIE is an automated genome project investigation environment; and PEDANT provides a ‘complete’ functional and structural characterisation of protein sequences. It is interesting, therefore, to examine more closely the tools and resources they each use. Thus, we find:

* GeneQuiz - uses BLAST/FastA, PROSITE, Blocks
* MAGPIE - uses BLAST/FastA, PROSITE, Blocks
* PEDANT - uses BLAST/FastA, PROSITE, Blocks

Over time, these systems have become more comprehensive, and today use additional databases, but even so, none is likely to give the complete structural and functional characterisation advertised by PEDANT, and none will offer reliable answers in the Twilight Zone of sequence analysis.

Other integrated approaches

Rather than build integrated systems at the level of the software tools, an alternative approach is to integrate the databases themselves - this is the principle behind the European InterPro project. InterPro attempts to simplify sequence analysis by amalgamating the family databases (PROSITE, PRINTS, Pfam, Profiles and ProDom) into a unified annotation resource, in which virtually all the documentation is provided by PRINTS and PROSITE - release 4.0 contains 4691 entries and as such is the most comprehensive family database worldwide. Additional partners are now joining the consortium, including TIGRfam, Blocks etc.. Although still relatively young, InterPro played a significant role in the annotation of the fly and human genomes.

CONCLUSIONS

When applying computers to the analysis of biological systems, it is important to understand what they can and can't achieve, and vital to appreciate the difference between mathematical and biological significance: computers don't do biology - they do sums....quickly! In light of this, it is clear that the success of functional analysis protocols based, for example, only on BLAST and PROSITE is likely to be limited.

Today, no single database is best. In any analysis, therefore, for best results, several tools and resources should be used, because different methods provide different perspectives, and because the databases aren't complete and their contents don't fully overlap. However, the more databases searched, the harder it can be to interpret results; hence more integrated and 'expert' systems are being developed to help users in the analysis process. Nevertheless, the more computers are used in sequence analysis, the greater the need for collaboration between software developers, database annotators and bench biologists. We are still far from having completely reliable in silico tools for functional characterisation of sequence data, but we have come a long way; and, with a concerted, considered approach, we can look forward to a promising future.

READINGS

1. Attwood, T.K., Blythe, M., Flower, D.R., Gaulton, A., Mabey, J.E., Maudling, N., McGregor, L., Mitchell, A., Moulton, G., Paine, K. and Scordis, P. (2002) PRINTS and PRINTS-S shed light on protein ancestry. Nucleic Acids Res., 30(1), 239-241.
2. Attwood, T.K. (2001) A compendium of specific motifs for diagnosing GPCR subtypes. Trends Pharmacol. Sci., 22(4), 162-165.
3. Attwood, T.K. (2000) The Babel of Bioinformatics. Science, 290, 471-473.
4. Attwood, T.K. (2000) The quest to deduce protein function from sequence: the role of pattern databases. Int. J. Biochem. Cell Biol., 32(2), 139-155.
5. Attwood, T.K. and Parry-Smith, D.J. (1999) Introduction to bioinformatics. Addison Wesley Longman, Harlow, Essex CM20 2JE, UK.
6. Henikoff, S., Greene, E.A., Pietrokovski, S., Attwood, T.K., Bork, P. and Hood, L. (1997) Gene families: the taxonomy of protein paralogs and chimeras. Science, 278, 609-614.

Gene Expression in Spermatogenesis using Microarray Technology

Paul J. Turek M.D.
Associate Professor of Urology, Obstetrics, Gynecology and Reproductive Sciences
University of California San Francisco, San Francisco, California, USA

Objectives. The participant should be able to:

- Describe the theory behind microarray application in the study of genetics
- Delineate the ways in which microarrays can be used to study spermatogenesis
- Define 3 different variables that influence microarray interpretation
- Understand how microarrays can be used for a) genetic class prediction and b) gene discovery

I. Introduction

In many organisms, at least 200-500 specific genes are required to make sperm. In humans, only a handful of loci and genes have been identified. Knowledge of such genes would likely further explain many cases of idiopathic infertility. Current clinical testing in genetic infertility has concentrated on examination of Y chromosome genes that cause spermatogenic failure. However, there is abundant evidence to suggest that the autosomes, in addition to the sex chromosomes, are equal or more important to male genetic infertility. Indeed, the vast majority of syndromes and conditions known to be associated with male infertility have been mapped to autosomal loci. In addition, the recent discovery of several distinct autosomal genes associated with infertility, including DAZL (DAZ-like) on chromosome 3p, and a human homologue of the fly sterility gene BOULE, designated hBOULE or BOL, located on chromosome 2q, support this contention (3,5). Thus, in the future, we should expect to find that hundreds of genes, in a wide variety of genomic locations, will be involved in infertility.

II. Microarray Technology

How are we going to find these genes and determine their importance in spermatogenic failure? Currently, pedigree analysis, chromosomal mapping, and expression analysis are widely used to help link genes to disorders. Among these, expression analysis offers the most potential for accomplishing our goal. Recent technical advances, largely derived from the computer chip industry, now make it practical to examine and quantify the expression of thousands of genes in parallel using complementary DNA (cDNA) microarrays (4). DNA microarrays are simply PCR-blot techniques applied on a massive scale. Gene sequences of interest are attached to glass slides at high density (thousands/cm). This array is then incubated with patient samples labeled with a color probe. If the patient's particular gene sequence is normal, then binding to the identical gene sequence on the array occurs and is detected by a color change (Figure 1).

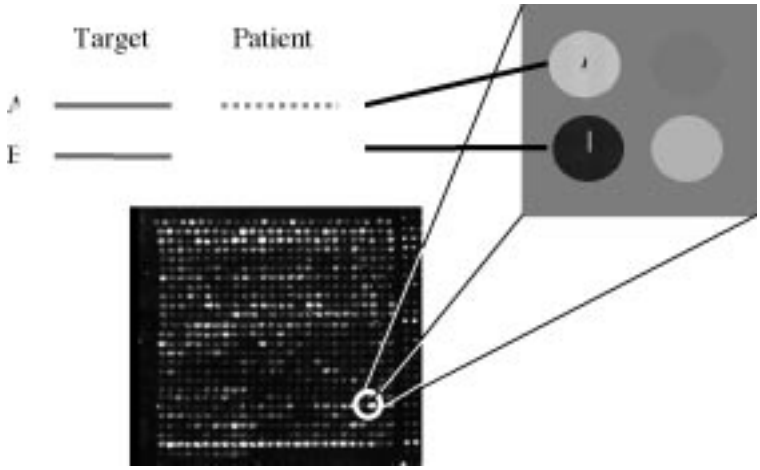


Figure 1. Analysis of microarray blots to determine gene expression

If the patient gene has a mismatched sequence, indicating a mutation, then hybridization efficiency is reduced or absent for that gene sequence. Subtle changes in hybridization efficiency can be detected with powerful software packages available for arrays (see www.microarrays.org). This incredibly powerful technology can be tailored to the individual biological system of interest and used to perform simultaneous genomic screening for thousands of mutations or examine patterns of gene expression (using RNA instead of DNA). This mode of genetic analysis is currently being used to observe variations in gene expression in a variety of human tumors, including colon, prostate, breast and ovarian cancers.

III. Practical Use of Microarrays

Despite their enormous potential, microarray technology is also loaded with pitfalls. The two largest problems with this technology relate to variability and statistical analysis. Variability occurs at 3 levels, each of which need to be understood to make sense of any data obtained:

- Platform or technical variability
- Variability in sample processing
- True biological variability

Only if the first two variables are well controlled, will the third have any meaning. In addition, upon viewing an array, it becomes immediately obvious that to apply normal paradigms of statistical analysis to this data is inconceivable. Instead, new and complex statistical methods have been developed to analyze array data, including cluster analysis, hierarchical clustering, and bootstrap analysis (2). The theory that underlies these analyses is the concept that genes that show similar expression patterns are likely to be related to each other. Thus, only with an understanding of these variables does the true value of array technology and the data it generates become apparent.

IV. Microarrays and Testis Genes

Our approach to microarray application in the testis is to use it as a tool for novel gene discovery. Early on, we sought to find testis genes relevant for meiosis. Initially, we created 3000 clones from a subtracted library generated by removing common transcripts between two different samples: pre-meiotic fetal testis and infertile testis tissue without germ cells (sertoli cell only). We hoped to obtain a library of unique or over-expressed genes likely relevant for spermatogenesis. These genes were placed on an in-house cDNA array and hybridized with human tissues and testis biopsies from obstructed and infertile men to identify transcripts relevant to testis biology. These experiments identified both housekeeping genes and potentially novel genes that require further elucidation (1).
A second approach seeks to use microarray-generated genes to improve on our current, morphologically based analysis of testis biopsies. Genetic analysis of testis biopsies holds far more potential than the currently outdated approach using microscopy in determining whether or not spermatogenesis is normal, and whether or not sperm are to be found within the failing testis. This form of testis genetic “class prediction” or “genetic signature” based on spermatogenic stage in fact represents one of the most powerful applications of microarray technology.

V. References

1. Ares X, Turek PJ, Reijo Pera RA. The use of cDNA microarrays to identify novel genes required for human male germ cell development. Fertil Steril. 76: P446A, 2001.

2. Eisen MB, Spellman PT, Brown PO and Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc. Nat. Acad. Sci. (USA).. 95: 14863-14868, 1998.

3. Saxena R, Brown LG, Hawkins T, et al: The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. Nat Genet 14: 292-9, 1996

4. Schena M, Shalon D, Davis RW and Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 270: 467-470, 1995

5. Xu EY, Moore FL, Reijo Pera RA: A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in metazoans. Proc Nat Acad Sci (USA) 98: 7414-19, 2001

Sequencing the Y - what it means

Dr. Renee Reijo Pera PhD.
Univ. of California at San Francisco
513 Parnassus Avenue, HSW 1480
San Francisco, CA 94143-0546
U.S.A.
Tel +1 (415) 476 179
Fax +1 (415) 476 31 21
E-mail reijo@itsa.ucsf.edu

The use of functional genomics in male infertility

1. The Y chromosome and male infertility.

2. The function of the *DAZ (Deleted in Azoospermia)* genes on the Y chromosome.

3. Strategies to identify new infertility genes.

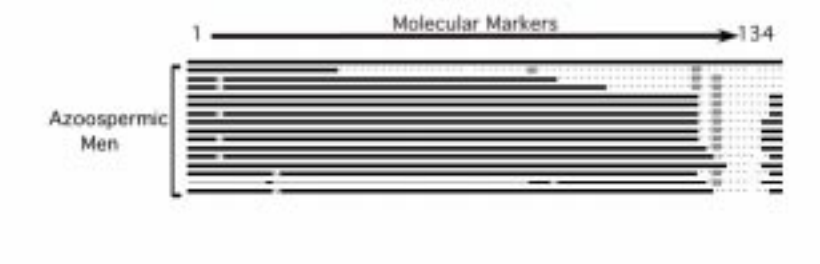
The Genetics of Human Infertility and Germ Cell Development

Infertility in humans, as in other organisms, may be genetic. It is likely that approximately 1000-5000 human genes may Function specifically in reproduction.
Genetic causes of reproductive failure may be linked to:

1. Genes required for somatic function such as formation of a functional reproductive tract, hormonal control of reproduction, or genes involved in the immune response.

2. Genes required for germ cell growth and development such as the genes required for oocyte maturation or sperm formation.

Overlapping Deletions in Azoospermic Men
but Not in Fertile Men

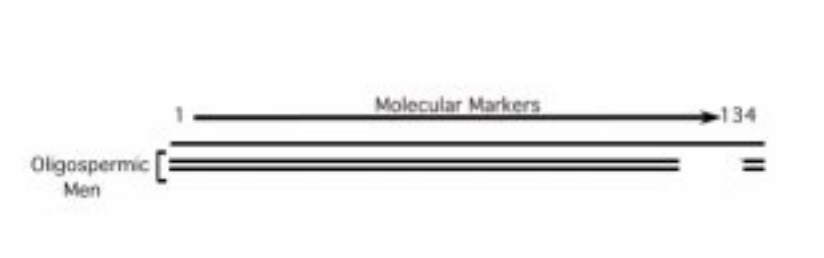


Is Oligospermia Caused by Deletions of the Y Chromosome?

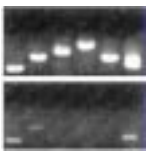
-Collected DNA samples from 32 men with sperm counts < 1 million/ml

-Assayed presence/absence of 134 PCR markers

Severe oligospermia resulting from deletions of the Azoospermia Factor gene on the Y chromosome



Sperm carry the deletions



ESHRE 2002 - Vienna

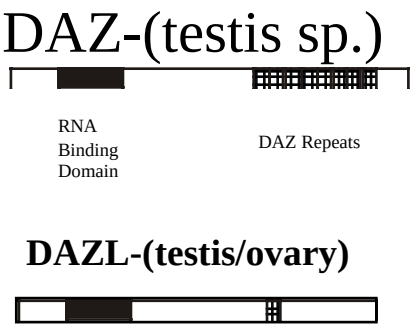
Hum Reprod 1999 Jul;14(7):1722-6

Men with infertility caused by AZF deletion can produce sons by intracytoplasmic sperm injection, but are likely to transmit the deletion and infertility.

Page DC, Silber S, Brown LG

Identification of the *DAZ* gene

MSTANPETPNSTISREASTQSSSA
ATSQGYILPEGKIMPNTVFVGGID
VRMDETEIRSFFARYGSVKEVKII
TDRTGVSKGYGFVSFFNDVDVQKI
VESQINFHGKKLKLGPAIRKQNLC
AYHVQPRPLVFNHPPPPQFQNVWT
NPNTETYMQPTTTMNPITQYVQAY
PTYPNSPVQVITGYQLPVYNYQMP
PQWPVGEQRSYVVPPAYSAVNYHC
NEVDPGAEVVPNECSVHEATPPSG
NGPQKKSVDRSIQTVVSCLFNPEN
RLRNSVVTQDDYFKDKRVHHFRRS
RAMLKSV



Evidence that deletion of the *DAZ* gene causes infertility

1. Men with deletions of the AZF region have deletions of DAZ; their fathers do not.

2. Studies in model organisms.

1. Meiotic cell cycle requirement for a fly homologue of human *Deleted in Azoospermia*

Eberhart CG; Maines JZ; Wasserman SA (Nature, 1996 Jun, 381:6585, 783-5).

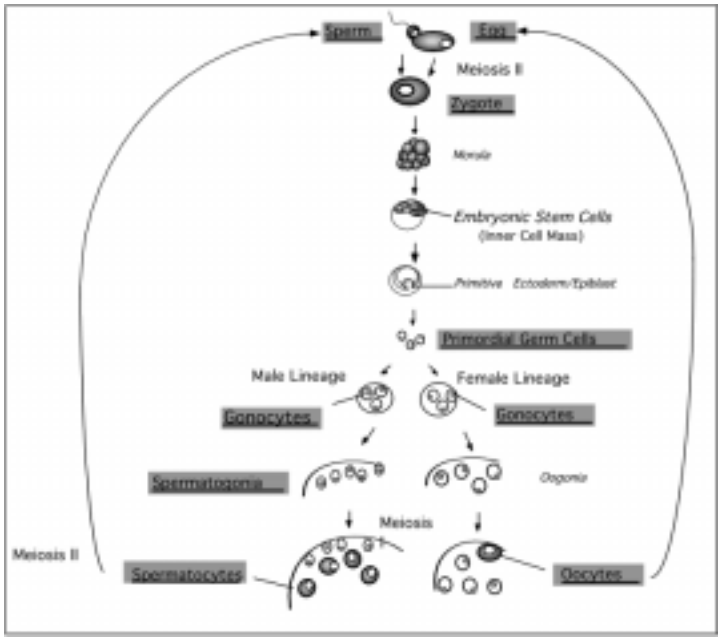
2. *Caenorhabditis elegans* homologue of the human azoospermia factor *DAZ* is required for oogenesis but not spermatogenesis.

Karashima T; Sugimoto A; Yamamoto M (Development, 2000 Mar, 127:5, 1069-79).

3. Mouse *Dazl* encodes a cytoplasmic protein essential for gametogenesis.

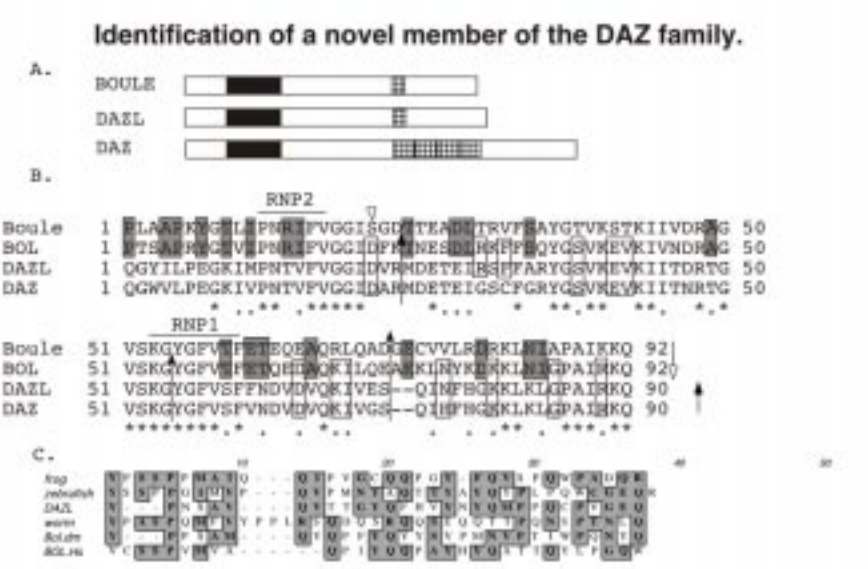
Ruggiu M; Speed R; Taggart M; McKay SJ; Kilanowski F; Saunders P; Dorin J; Cooke HJ (Nature, 1997 Sep, 389:6646, 73-7).

Expression of *DAZ* and *DAZL*

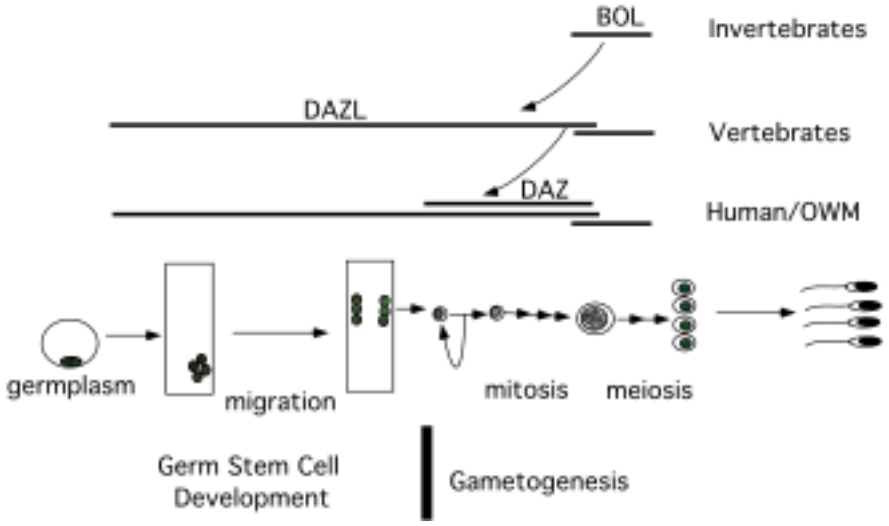


POTENTIAL INTERACTING PROTEINS WITH DAZ(L)

PUMILIO
hBOULE
DZIP (ZINC FINGER PROTEIN)
QUAKING PROTEIN



Acquisition of Novel Functions for the DAZ(L) Genes



What causes infertility in men without Y chromosome deletions?

Hints:

1. Model Organisms - DNA repair and recombination defects.

2. A single report in the literature.

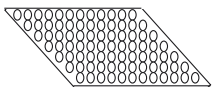
Cytogenetics 1970;9(6):460-7

Pearson PL, Ellis JD, Evans HJ

A gross reduction in chiasma formation during meiotic prophase and a defective DNA repair mechanism associated with a case of human male infertility.

Determine whether meiotic arrest in men who present with very low sperm counts or azoospermia is caused by defects in the recombination and/or DNA repair pathways.

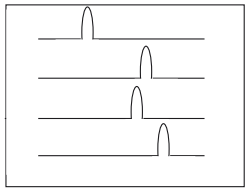
DNA Repair



Sort Sperm

Prepare DNA

Genotype



Recombination

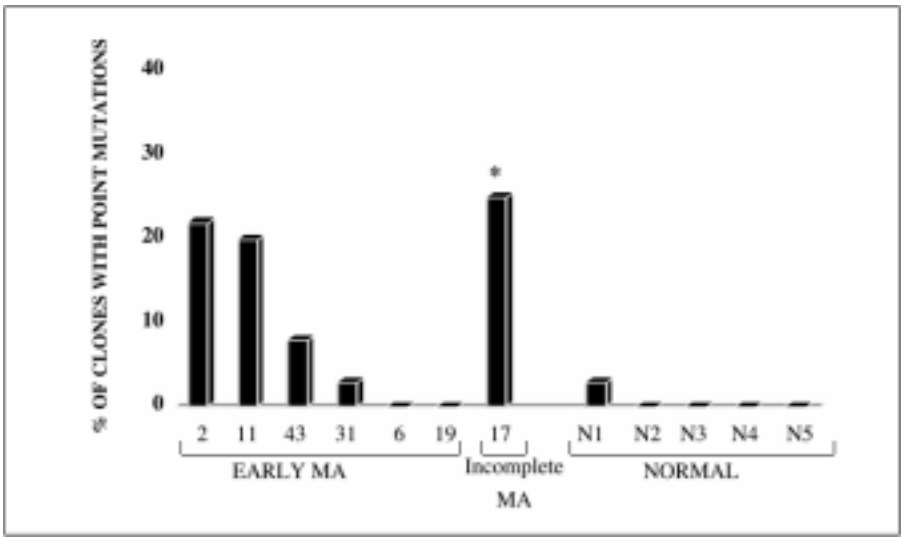
Maternal

Paternal

Determine % Recombinants

Compare % Recombination in Infertile vs. Fertile Men

DNA Repair in Infertile Men



Parameters for Reliable Results in Genetic Association Studies

- plausible biological context
- low *P* values
- independent replication
- rigorous phenotypic assessment
- rigorous genotyping
- appropriate statistical analysis
- alleles affect the gene product in a physiologically - meaningful way

KO mice - from mouse to man, what can we learn

Edward Mitchell Eddy, Ph.D., Na
tional Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709 USA

Upon completion of this presentation, participants will know:

* How and why a gene knockout (KO) is done
* KOs can disrupt male reproduction directly or indirectly
* A wide variety of novel genes are required for spermatogenesis
* KOs provide novel insights into male reproductive function

HOW AND WHY A GENE KNOCKOUT (KO) IS DONE

The development of gene knockout (KO) technology has resulted in substantial advancements in knowledge in many areas of mammalian biology. In the last decade it was shown that: 1) mutations can be targeted to specific genes in embryonic stem (ES) cells by homologous recombination, 2) ES cells in which this rare event has occurred can be cloned using positive and negative drug selection strategies, and 3) ES cells containing the mutant gene can be used to produce chimeric mice that give rise to progeny bearing the mutation in every cell. Other essential methods developed previously made this technology possible. They include isolating ES cells from the inner cell mass of preimplantation mouse embryos, maintaining ES cells in culture without undergoing differentiation, and producing chimeric mice by injecting ES cells into mouse blastocysts. So far, totipotent ES cells have been isolated only from mice. The production of gene KO mice requires skills in molecular biology, tissue culture, and micromanipulation of mouse embryos, as well as knowledge of mouse reproductive biology and colony management. However, the greatest challenge often occurs after the knockout mouse is produced, that of determining how the mutation has led to the phenotype found.

Transgenic mice

The term “transgenic mice” is sometimes used to describe mice whose genome was modified by either of two approaches. One is to introduce an artificial gene that integrates randomly, and the other is to use homologous recombination to modify an endogenous gene. However, the most common use of the term transgenic mice refers to those made by injecting an artificial gene into the pronucleus of a zygote. The artificial gene or transgene usually consists of the promoter from one gene ligated to the protein-coding region of another gene. These are sometimes called “promoter-reporter constructs.” The promoter determines where and when the transgene is expressed. This expression may be ubiquitous, restricted to a particular tissue or organ, initiated at a specific time during development, or in response to treatment with a drug or another agent. The mouse produced from an injected zygote is referred to as the founder of a transgenic line. Not all transgenic lines that contain the same transgene are equivalent. Some have silent transgenes that fail to express because the transgene integrated into a region of the genome where gene expression is suppressed (e.g., ends of chromosomes, near centromeres). Other lines have promoter-trapped transgenes that are integrated immediately downstream to the promoter of an endogenous gene, resulting in their expression being determined by the endogenous promoter. In addition, there are often

differences between lines in the level of transgene expression. This is probably caused by positional effects due to proximity to other genes, formation of concatemers of the transgene during integration, or integration of transgenes at multiple sites in the genome. Transgene can also result in insertional mutagenesis by integrating into the promoter or the protein-coding region of an endogenous gene and disrupting its expression.

Gene knockout mice

ES cells are derived from totipotent cells of blastocyst-stage embryos. Intact embryos are placed in culture, cells grow out of the inner cell mass in a few days, and these cells are isolated and clones are maintained in vitro on feeder layers of embryonic fibroblasts. A gene-targeting construct usually consists of two portions of a gene that are separated by a positive selection gene and flanked by a negative drug-selection gene. The construct is electroporated into ES cells and if homologous recombination occurs, the negative selection gene is deleted and the positive selection gene (neomycin resistance) renders the cell resistant to G418. However, if integration occurs without homologous recombination, the negative selection gene (herpes simplex thymidine kinase) is retained and the cell is killed by gancyclovir. Homologous recombination is a rare event in transfected ES cells, but positive and negative selection greatly improve the probability of producing ES cell clones with successfully targeted genes. Similar approaches may be used to replace a gene with a different gene, to duplicate a gene, or to precisely introduce small mutations in a gene.

Gene knockout mice are made by injecting ES cells into the cavity of a day-3.5 blastocyst and then putting the blastocyst into the uterus of a surrogate mouse to allow development to continue. Mice produced are chimeras containing a mixture of cells derived from the inner cell mass of the blastocyst and from ES cells. The chimeras are mated and offspring are produced with the modified gene in their genome. These methods are described in detail in Hogan et al. (1).

Applications

Genome sequencing has shown that the similarities between genes in humans and mice (and many other animals) enormously out-weight the differences. While the mouse is not a perfect surrogate in which to study genes with essential roles in human health, KOs have been used quite successfully as models of many heritable human diseases (2). These include diseases caused by point mutations (e.g., lysosomal storage diseases), deletions (e.g., DiGeorge syndrome) and diseases with multigenic causes (e.g., diabetes). KO mice are being used to determine where and when the effects of a mutation is first detected, to define the progression of disease processes, to develop new diagnostic tests, and to evaluate the effectiveness of drug treatments and gene replacement therapies. In addition, KO mice are used to gain fundamental knowledge about the regulation of genes and the function the proteins they encoded, within the context of the intact animal. Most of the genes targeted in KO studies are compiled in the Mouse Knockout Mutation Database (3). KOs are particularly valuable for studying the roles of specific proteins in organ systems and of complex tissues that are not accessible with in vitro systems.

KOs CAN DISRUPT MALE REPRODUCTION DIRECTLY OR INDIRECTLY

Spermatogenesis provides a particularly good example of how the use of KOs has led to new insights about the regulation and function of essential cellular, tissue, organ, and organ system processes. The effects of KOs on male reproduction have ranged from a modest reduction in fertility with no obvious changes in spermatogenesis to sterility due to the absence of the testis. Some KOs with effects on male fertility disturb development processes, while others alter endocrine processes or modify the function of somatic cells in the testis. In some cases the disrupted gene is expressed only in somatic cells of the testis (e.g., Rxrb), while in others it is expressed only in spermatogenic cells. Some KOs expected to alter male fertility have little or no effect (e.g., Acr), while others unexpectedly disrupt spermatogenesis or processes essential for normal male fertility (Pms2).
Gene KOs produce diverse effects on spermatogenesis because of how gene expression is regulated in these cells. This occurs at three levels, intrinsic, interactive, and extrinsic. A highly conserved genetic program intrinsic to germ cells determines the sequence of events that underlie germ cell development. The intrinsic program determines which genes are utilized during processes specific to germ cells and when they are expressed (e.g., meiosis, haploid gene expression, translational delay, expression of novel genes and transcripts). However, spermatogenesis is not a cell autonomous process. Male germ cells develop in overlapping waves, with cohorts of germ cells developing in synchrony. The intrinsic program operating within a particular germ cell requires information from, and provides information to, neighboring cells to achieve this coordination. Sertoli cells are crucial at this level of interactive regulation, as well as for providing essential support for germ cell proliferation and progression through the phases of development. The interactive level of regulation is in turn dependent on extrinsic influences, primarily steroid and peptide hormones, which act through Sertoli cells and peritubular cells to impart indirect effects on germ cells. The extrinsic and interactive processes are indispensable for establishing and maintaining an optimum environment within which spermatogenesis occurs.

KOs can disrupt male reproduction directly
Germ cell development occurs in successive mitotic, meiotic and post-meiotic phases during which the germ cells move from the periphery to the lumen of the seminiferous tubule. The mitotic phase takes place in the basal compartment, while the meiotic and post-meiotic phases occur in the luminal compartment. This process takes about 35 days in the mouse, with the mitotic phase lasting approximately 11 days, the meiotic phase lasting approximately 10 days, and the post-meiotic phase lasting approximately 14 days. The process begins every 8.7-8.9 days and the duration of these phases varies slightly between mouse strains. It includes processes unique to germ cells, including meiosis, genetic recombination, haploid gene expression, formation of the acrosome and flagellum and remodeling and condensation of chromatin, all of which must occur at the correct time to produce the male gamete.
Mitotic phase: During the mitotic phase, stem cells give rise to spermatogonia, which divide six times in the mouse. However, approximately half of the spermatogonia undergo apoptosis and members of the Bcl2 gene family have critical roles in this process. Bax promotes apoptosis and the increase in number of spermatogonia in Bax KO mice leads to disruption of spermatogenesis and infertility. Transgenic mice over-expressing BclxL, an apoptosis-inhibiting protein, also have excessive numbers of spermatogonia and are infertile.
Meiotic phase: The last cell cycle S phase occurs at the end of the mitotic phase, after which the unusually long 8-day long G2 phase of prophase I of meiosis begins. The novel process of meiosis involves pairing of homologous chromosomes and chromosomal recombination, and ends with desynapsis and two reductional cell divisions, producing four haploid spermatids from each spermatocyte. Disruption of these processes by gene KOs usually results in failure of spermatogenesis. For example, the synaptonemal complex (SC) is the structure that binds together homologous chromosomes along their

length, and the KO of the gene for synaptonemal complex protein 3 (Sycp3), disrupts SC assembly and chromosome synapsis, causing male infertility. In addition, recombination involves breaking and repairing of the DNA strands. The KO of genes for components of the DNA repair machinery (e.g., Pms2, Dmc1h, Mlh1) often results in male infertility, while producing no overt effects in other cells. Many of the genes involved in these processes are reviewed in the volume edited by Handel (4).
Post-meiotic phase: Although spermatids have a haploid genome, there is a high level of gene transcription during the first part of this phase. However, the translation of a number of the mRNAs produced during this period is delayed for several days. In addition to these novel processes, spermatids develop an acrosome and a flagellum, and replace the nuclear histones with transition proteins and those in turn with protamines, to bring about compaction of the chromatin. KOs of genes for most of the proteins involved in these processes disrupts male fertility. For example, the CREM(transcription factor regulates expression of several genes expressed during the post-meiotic phase and the KO of its gene disrupts early spermatid development, causing male infertility. Furthermore, disruption of one allele of the gene for protamine-1 or -2 results in haplo-sufficiengy, the reduction in amount of protein, resulting in failure of chromatin compaction and causing male infertility.

KOs can disrupt male reproduction indirectly
KOs can disrupt male reproduction indirectly by perturbing development, endocrine function, or Sertoli cell function. Examples of KOs that disrupt male reproduction by having detrimental effects on development of the testis include genes for homeobox transcription factors (e.g., Hoxa10, Hoxa11, Sp4), anti-Mullerian hormone (Amh), a member of the TGF-beta superfamily of growth factors, and the anti-Mullerian hormone receptor (Amhr2). Examples of KOs that disrupt male reproduction by disrupting endocrine function include inhibin-alpha (Inha) and its receptor (Acvr2a), involved in gonadal-hypophyseal feedback, and estrogen receptor alpha (Estra). Examples of KOs that disrupt Sertoli cell function include genes for a cell cycle regulatory protein, cyclin D2 (Cycd2), transcription factors (e.g., Dhh, Fmr1), and a receptor retenoid (Rxrb).
In vitro studies have shown that the relationship between germ cells and Sertoli cells is bi-directional. Sertoli cells are essential for germ cells to proliferate and to progress through their phases of development. The general functions of Sertoli cells in these processes are well known, such as dividing the seminiferous epithelium into basal and luminal compartments and providing a suitable milieu within which germ cell development occurs. In addition, the patterns of gene expression in Sertoli cells change as neighboring germ cells progress through the stages of spermatogenesis. This suggests that diverse signals coming from the cohort of developing germ cells act on Sertoli cells and cause changes in gene expression, and this may in turn alter feedback signals that modulate the intrinsic regulatory program driving germ cell development. For example, KO studies indicate that glial cell line-derived neurotrophic factor (GDNF) produced by Sertoli cells signals germ cells.

A WIDE VARIETY OF NOVEL GENES ARE REQUIRED FOR SPERMATOGENESIS

The expression of many germ cell-specific transcripts is developmentally regulated and stage specific. Some of these transcripts are product of genes that are male germ cell-specific homologs of genes expressed in somatic cells, while others are expressed from unique genes unlike any others in the genome. In addition, there are many alternate transcripts that are products of genes also expressed in somatic cells, but differing from them in size and/or overall sequence. These transcripts are produced by germ cells using promoters and transcription factors that activate transcription at different start sites upstream or downstream of the usual site, by incorporation of alternate exons, by germ cell-specific splicing events,

and by using alternate initiation sites for polyadenylation. Male germ cell development consists of an assortment of unique processes that are intricate and highly ordered, and require novel gene products and a precise and well-coordinated program of gene expression for them to take place. Many of the genes involved in these processes are included in an on-line Mouse Reproductive Genetics Database (5).

Male germ cell-specific homologs

Genes expressed only in male germ cells frequently are homologs of genes expressed in somatic cells (reviewed in Eddy and O'Brien, 1998). In some cases, a gene that is expressed in somatic cells is inactivated and a germ cell homolog is inactivated. An example of this is the spermatogenic cell-specific form of glyceraldehyde 3-phosphate dehydrogenase (Gapds), the homolog of a highly conserved gene for an essential enzyme in the glycolytic pathway. Gapd is usually considered to be a housekeeping gene and its mRNA is commonly used as a loading control on northern blots. However, the Gapd gene is inactivated during the meiotic phase and the Gapds gene is activated early in the post-meiotic phase of male germ cell development.

It seems inefficient for male germ cells to use a new gene instead of an existing gene that encodes a nearly identical protein. However, there are often significant advantages for male germ cells to have their "own" genes. One is that a germ cell homolog may compensate for a gene that is inactivated in male germ cells. An example of this is the phosphoglycerate kinase-1 (Pgk1) gene that encodes another essential enzyme in the glycolytic pathway. Pgk1 is located on the X chromosome, which is inactivated during the meiotic phase of male germ cell development. However, the germ cell-specific Pgk2 gene is located on an autosome and is activated soon thereafter. Another advantage of germ cell-specific genes is that the encoded protein may fulfill the same role as that of somatic cells product, but also have other structural or functional properties that serve unique roles in male germ cells. For example, GAPDS is anchored to the fibrous sheath by a proline-rich N-terminal domain that is absent in GAPD.

Activation of germ cell-specific gene expression may also occur without the associated inactivation of a homolog. An example of this occurs for a member of the 70-kDa heat-shock protein (HSP70) family. The HSP70 proteins are chaperones that assist other proteins to fold as they emerge from the ribosome, escort them through the cytoplasm, help them assemble into complexes, and facilitate the refolding of proteins partially denatured by heat or other stresses. The genes for most HSP70 proteins are expressed in all tissues either constitutively or following induction by heat shock and other stresses. However, a major 70 kDa protein synthesized in spermatocytes is a novel member of the HSP70 family (Hsp70-2, and male but not female mice with a KO of the gene are infertile. Germ cells in the mitotic and meiotic phases are present in the testis, but development is arrested at the late pachytene spermatocyte stage and these cells undergo apoptosis. The HSP70-2 protein apparently serves a chaperone role that is specifically required in male germ cells that is not provided by the other HSP70 proteins present.

Unique genes

Some genes expressed only in male germ cells do not have homologs in the genome. Most of these genes are expressed during the post-meiotic phase when the specialized structural components of spermatozoa are produced. The transition proteins and protamines mentioned previously are products of unique genes expressed only in male germ cells. Other examples are proteins found only in the fibrous sheath of the flagellum, AKAP3 and AKAP4.

Alternate transcripts

Male germ cell-specific proteins may also be synthesized from alternate transcripts that are derived from the same gene as transcripts in somatic cells, but differ from them in size and/or overall sequence. They are generated during gene expression by the use of promoters and transcription factors that activate transcription at different start sites upstream or downstream of the somatic cell site, by incorporation of alternate exons, by germ cell-specific splicing events, and by using alternate initiation sites for polyadenylation. Transcription initiation sites within introns are used to produce truncated versions of angiotensin-converting enzyme (ACE) and calcium2/calmodulin-dependent kinase IV (CaMKIIV) that are found only in male germ cells and have enzymatic properties different from the longer forms.

KOs PROVIDE NOVEL INSIGHTS INTO MALE REPRODUCTIVE FUNCTION

Gene KO have identified diverse genes that are essential for male reproduction, revealed that some aspects of male endocrinology are still not well understood, and demonstrated that genes expressed specifically in germ cells often have unique roles in male reproduction. The information gained from KO studies has provided a better understanding of the processes and mechanisms involved in male germ cell development and function and of the genes required for male fertility.
It is commonly believed that FSH and testosterone are the regulators of spermatogenesis. However, it has become increasingly apparent that FSH and testosterone do not regulate germ cell development, but instead serve an important supportive role for this process. FSH is not essential for spermatogenesis, as show by the finding that male mice with a KO of the gene for FSH-beta are fertile and that men with an inactivating mutation in the FSH receptor have varying degrees of spermatogenic failure, but are fertile earlier. In addition, testis weight and epididymal sperm numbers are reduced in FSH receptor KO mice, but the males are fertile. It has become apparent that germ cell development is considerably less dependent on FSH than was previously thought.
The other major player in the extrinsic regulation of germ cell development, testosterone, is essential for this process. LH regulates the production of testosterone by Leydig cells and testosterone acts through androgen receptors in Sertoli cells and peritubular cells to influence the Sertoli cell functions that support germ cell development. In the absence of LH, spermatogenesis can be maintained by administering testosterone, but not at the same level as in the normal state. Although it has been reported by some that androgen receptors are present in germ cells, it has been shown recently that they are not essential for germ cell development. Using the germ cell transplantation technique, germ cells from the testis of Tfm mice lacking functional androgen receptors were transplanted into the testes of mice with functional androgen receptors. Colonies of donor-derived spermatogenic cells were seen in the seminiferous tubules of recipient mice, indicating that mouse germ cells do not require functional androgen receptors to carry out spermatogenesis.

Lessons learned from KO mice: Male fertility

Endocrine function

Genes required: estrogen receptor-alpha (Estr1), inhibin-alpha (Inha)
Genes not required: estrogen receptor-alpha (Estr2), FSH-beta (Fshb), FSH receptor (Fshr), androgen receptor (Ar) in germ cells

Development

Genes required: anti-Mullerian hormone (Amh) and anti-Mullerian hormone receptor (Amhr), c-ros oncogene, multiple transcription factors

Spermatogenesis

Genes required: homologs of genes expressed in somatic cells, novel genes for germ cell-specific components, germ-cell specific alternate transcripts, apoptosis genes in Bcl2 family (Bax, Bclw), DNA repair genes

Sperm function

Genes required: protamine-1 and -2 (Prm1, Prm2)

Genes not required: acrosin (Acr), transition protein-1 and -2 (Trp1, Trp2), histone H1t (H1t)

Human syndromes similar to KO phenotypes

Hermaphroditism - Amh, Amhr
Cryptorchidism - Hoxa10, Hoxa11, Insl3
Incomplete epididymis - c-ros
Sertoli-cell only - c-kit receptor and ligand
Tubular degeneration - Bmp8b
Meiotic arrest - Hsp70-2, Cyca1, A-myb, Dmc1, Atm, Mlh1
Post-meiotic arrest - Cremt, Tl,
Teratospermia - Akap4, Hr6b, Pms2, Prm1, Prm2
Reduced or no fertility, normal sperm numbers and motility - Sprm1, Psck4, Ftnb, Clgn, tACE
Reduced or no sperm motility - Dnahc1, Catsper

References

1). Hogan, B., Beddington, R., Costantini, F., and Lacy, E. (1994) Manipulating the Mouse Embryo. 2nd edn. Cold Spring Harbor Laboratory Press. Cold Spring Harbor.

2. Online Mendelian Inheritance in Man. (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=OMIM)
This database is a catalog of human genes and genetic disorders. It contains textual information and references. It also contains links to MEDLINE and sequence records in Entrez, and links to additional related resources at NCBI and elsewhere.

3. Mouse Knockout and Mutation Database. (MKMD; http://research.bmn.com).
A comprehensive on-line database that is a useful source of references for mouse knockout studies is available by subscription. It also contains a review section on Mouse Models of Human Disease.

4. Handel, M. A. (1998) Meiosis and Gametogenesis. Current Topics in Developmental Biology, Vol. 37, Academic Press, San Diego.
Chapters by various authors in this volume provide reviews on genetic recombination, DNA repair, the genetic control of meiosis, and genes expressed during the meiotic phase.

5. Mouse Reproductive Genetics Database. (http://mouse.genetics.washington.edu/).
Information on genes and literature related to male reproduction is contained in this database. You can search the database for genes and literature based on phenotypes, expression profiles, and function.

Gene causing defects in spermatogenesis

David M. de Kretser, MD
Monash Institute of Reproduction and Development,
c/o Monash Medical Centre,
246 Clayton Road, Clayton,
Melbourne, Victoria,
Australia 3168

Learning Objectives

1. To understand the key physiological steps at which genetic defects could disrupt spermatogenesis and cause infertility.
2. To appreciate the approaches which have resulted in our current understanding of genetic defects that disrupt fertility in men.
3. To obtain an appreciation of the current state of knowledge of genetic mechanisms that can disrupt fertility.

Introduction

For the clinician, the inability to define the cause of a spermatogenic defect and develop an evidenced-based treatment regime, is one of the most frustrating problems. There is now an emerging, reasonably well developed body of evidence that genetic defects are causally related to a major proportion of the spermatogenic defects that currently are classified as “idiopathic spermatogenic disorders”. The earlier papers in this course have highlighted the rapidly accumulating data from other species concerning genetic mechanisms that can cause defects in spermatogenesis. In this paper, the known and emerging genetic mechanisms are assembled into a framework that can assist the clinician in the application of these developments to patient management and to facilitate clinical research into genetic mechanisms.

Physiological Framework

There are multiple control points in the physiology of spermatogenesis that could be disrupted by mutations in specific genes. The resultant defects are divisible into two major groups, those that are testis specific and result in infertility as the only phenotype and others that involve alterations in many organ systems. A clear example of the latter are the mutations in the cystic fibrosis transmembrane regulator (CFTR) which cause cystic fibrosis and congenital absence of the vas deferens and yet in some instances cause predominantly agenesis of the vas (Chillon et al, 1995).

Key processes in male reproduction that can be disrupted may be classified as follows:

I. Hormonal Mechanisms
II. Mechanisms Controlling Spermatogenic Output
(A) Migration of Germ Cells into Developing Gonad
(B) Spermatogonial Proliferation and Survival
(C) Meiosis
(D) Spermiogenesis
(E) Multiple Check Points
III. Leydig Cell Defects
IV. Sperm Transporting System

The Approach to the Identification of Genetic Defects

The approaches that have been undertaken to identify genetic defects causing infertility have carried. Frequently, these have arisen from the identification of a crucial control mechanism with a subsequent targeted disruption of a key gene leading to a phenotype in mice (see review Cram et al). This has been followed by a search for the equivalent human phenotype. Examples of this approach in identifying the human counterpart in men with infertility are rare as yet but will increase rapidly in the future. One example relates to our understanding of the need for Sertoli cells to produce stem cell factor that acts, through its receptor c-kit on spermatogonia, to stimulate spermatogonial mitosis and survival. This knowledge has been gained by studying naturally occurring mutations in the c-kit and stem cell factor genes(Loveland). A mutation in the c-kit gene in mice results in the white-spotted mutant where failure of normal melanocyte migration (creating the white spots), anaemia and infertility coexist, indicating crucial actions in all systems. A similar mutation in the c-kit gene in certain families results in the condition of human piebaldism, due to failure of normal melanocyte migration, but did not result in infertility or anaemia(ref).

To date most of the specific gene defects causing infertility have arisen from identification of a disorder with a clear familial transmission with a subsequent search for the specific gene defect. The elucidation of the genetic mechanism of androgen insensitivity syndrome arose from the identification of the hormonal mechanism of androgen insensitivity and the identification of the X chromosomal location of the androgen receptor gene. Subsequent identification of mutations in this gene demonstrated the genotype-phenotype linkage. While defective spermatogenesis was not the issue that gained clinical attention in the early reports of these patients, that were dominated by the male phenotype with female external genitalia, there was no doubt that spermatogenesis was disrupted. More recent studies have identified specific mutations in the androgen receptor gene that cause defective spermatogenesis and infertility without altering the male external genitalia. Further linkage of the androgen receptor gene and spermatogenic defects relates to the expanded CAG repeat (polyglutamine tract) in exon 1. Several studies have linked a high risk of azoospermia/severe oligospermia to expansions of the CAG repeat beyond 26 [mean for populations 20.7 (USA), 22.4 (Singapore), 20.8 (Australia), 21.8 (Denmark)]. These and other observations suggest that the androgen receptor gene with an expanded CAG repeat has low intrinsic androgen receptor activity (Tut et al, 1997; Dowsing et al, 1999; Mifsud et al, 2001; Rajpert-De Meyts et al, 2002). Others have argued against this linkage but often these studies aggregated all infertile men in their populations thereby potentially obscuring the specific mechanism. Those studies identifying the linkage evaluated men where other known causes of infertility including Y chromosome deletions were excluded.

A second area arising from observations from karyotype linked observations is the field of Y chromosome deletions, emerging from the original finding of small Y chromosomes associated with azoospermia(Tieopolo). Subsequent improvements in our ability to localize defects has culminated in the identification of the nature of the deletions and their frequency which is conservatively estimated to be between 6-10% of men with idiopathic seminiferous tubule failure and about 2% of all men with sperm counts less than 10 million/ml.

In a further example of such an approach, Olesen et al. ( ), used digital differential display to identify testis expressed transcripts and compared their chromosomal mapping position to the breakpoints found in men with balanced reciprocal translocations found in 265 infertile men. They identified several "hot spots" at 1p31-33, 6p21, 6p22.1, Xq28, 7q 31 and 3p21.1-9. Some of these foci represent regions where known testis expressed genes are located and others represent the sites of novel genes with respect to testicular function.

A third area demonstrates the extraordinary power of modern genomics. It has been recognised that the immotile cilia syndrome (now termed primary ciliary dyskinesia) was an inherited disorder. The phenotype of immotile cilia causing infertility and respiratory disorders such as bronchiectasia is associated with abnormalities of the axoneme where the dynein arms (a protein complex with ATP-ase activity) were absent on electron microscopy. Recognizing that the structure of cilia were highly conserved even to algal organisms such as chlamydomonas, recent studies have used mutant algal forms to identify some of the key genes. Subsequent examination of the human genome identified two genes DNAI1 and DNAH5 with very high sequence homology (approximately 90%) to the homologs in chlamydomonas (Pennarum et al, 1999; Olbrich et al, 2002), both of which show point mutations in patients with primary ciliary dyskinesia lacking outer dynein arms.

Finally, the identification of crucial control mechanisms have led to targeted disruption of a key gene leading to a phenotype in mice subsequent to which a search for the equivalent human phenotype has been undertaken. Examples of this approach up to now are sparse. Targeted disruption of the gene encoding the steroid acute regulatory protein (STAR) established a phenotype which affected all steroid producing glands such as the adrenal, testis and ovary producing a complex phenotype (Hasegawa et al, 2000).

It is surprising that more examples of this approach leading to elucidation mechanisms are not available given the ever increasing number of targets that arise from gene knock-out experiments in mice. In part this is no doubt due to the relatively small number of academic institutions focusing on andrology that can collect detailed clinical information and DNA samples. Our own approach has been to develop a DNA collection with accurate clinical information and quality data such as semen analysis, testicular histology and electron microscopic imaging of sperm tails (from men with motility disorders). This has led to a collection of over 2000 samples of DNA. Accurate histological reporting on testis biopsies when performed is crucial to defining the stage of spermatogenesis that is disrupted. The latter requires fixation in fluids such as Bouin's and not formalin which causes shrinkage and loss of chromatin patterns essential for identification of germ cell type.

Can We Gain More Understanding from Gene Knock-Out Models?

In our current quest to publish in high impact factor journals, many papers concerning gene knock-out phenotypes have limited data, due to constraints of space concerning specific details of testicular phenotypes especially if there are abnormalities in multiple systems. Further sometimes, detailed cytological analysis is not performed due to the time taken and number of animals needed especially if a developmental study is required.

For instance, the phenotype in the Bcl-w knock-out mice would have been inadequately characterized in detail if quantitative studies had not been performed demonstrating increased apoptosis prior to the profound "collapse" of spermatogenesis at six weeks and thereafter (Print et al, 1998). Follow up studies identified that in the absence of Bcl-w in the first 28 days (?) other pro-survival molecules were present but their expression declined in the adult testis leading to cell death and profound germ cell depletion (Meehan et al, 2001).

In other studies, it was noted that targeted disruption of the (-subunit of FSH led to disruption of folliculogenesis but maintenance of fertility in males albeit with lower testicular volumes (Kumar et al, 1997). More details morphometric studies have confirmed that the lower testis volume and decreased sperm output are in part due to a decrease in Sertoli cell number due to the absence of the proliferative action of FSH. However, these studies showed that the number of germ cells that could be supported by an individual Sertoli cell decreased indicating a metabolic requirement for FSH to maintain the "carrying capacity" of the Sertoli cells (Wreford et al, 2001).

Specific Defects in Causing Infertility in Humans and Mice

A number of earlier presentations have explored some of the approaches used to identify genes causing infertility in men. A detailed consideration of them in this paper is unwarranted within the time allowed. Rather, this paper has assembled many of the genetic defects disrupting fertility into a format that allows a clinical approach to the identification of known genetic defects in patients and facilitates confirmation of genetically determined spermatogenic defects in mice as relevant to man. The Tables 1-4 utilise the physiological framework identified earlier to categorize the knowledge to date in some logical arrangement.

Table 1
Hormonal Mechanisms

(Gonadotrophin releasing hormone
- Production : Kalig 1 gene

(FSH
- Mutations in (-subunit
- Mutations in (-subunit common to LH and TSH
- Mutations in FSH receptor

(LH
- Mutations in (-subunit
- Mutations in (-subunit common to FSH and TSH
- Mutations in LH receptor

(Testosterone
- Mutations in steroid biosynthetic enzymes, some common to adrenal
- Mutations in steroid acute regulatory protein gene
- Mutations in androgen receptor including CAG repeat

Table 2
Mechanisms Controlling Spermatogenic Output

A. Migration of Germ Cells into Developing Gonad
- C-kit (stem cell factor receptor) mutations
- Stem cell factor mutations
- RNA-binding protein TIAR KO

B. Spermatogonial Proliferation and Survival
- Bax - (proapoptotic)
- Apoptosis protease-activating factor (Apat-1)
- DFFRY
- AZFa deletions
- Dnmt 3L: ?spermatogonial loss; Sertoli cell only phenotype

C. Defects in Meiosis
- Bcl 16 (antiapoptotic): (apoptosis in M1
- Ataxia telangiectasia mutant (ATM): chromosome fragmentation
- Cyclin A1: Desynapsis abnormalities at M1
- Deleted in azoospermia-like (Dazla): loss in M1
- Dmc1-meiosis specific RecA: zygotene arrest
- HSP70.2: synaptonemal complex desynapsis failure
- MLH1-DNA mismatch repair enzyme: meiosis arrest
- RAD6b (hr6b)-ubiquitin conjugating enzyme: postmeiotic chromatin condensation
- Synaptonemal complex protein 3 (SCP3): chromosome synapse failure
- Translocated in liposarcoma (TLS): failure of synapsis
- Microorchidia (morc): zygotene-leptotene arrest
- Siah 1a: Failure of M1 metaphase to anaphase transition
- Mouse vasa homolog gene (Mvh): zygotene arrest
- AZFb deletions: meiosis arrest

D. Specific Defects in Spermiogenesis
- Casein kinase II catalytic subunit (CK2): globozoospermia
- Cyclic AMP-responsive element modulator (CREM): early spermatid arrest

- Ca++/calmodulin dependent protein kinase IV (Camk4): elongating spermatid defect
- c-ros (orphan tyrosine kinase receptor): sperm tail angulation and tangling
- Transition nuclear protein 1 (TP1): decreased sperm motility
- Apolipoprotein B (apo B): decreased sperm motility and survival
- DNAI1: loss of outer dynein arms in primary ciliary dyskinesia
- DNAI2: candidate for primary ciliary dyskinesia
- MDHC7 (mouse dynein heavy chain): KO (ciliary dyskinesia
- DNAH5: absence of outer dynein arms, primary ciliary dyskinesia
- Tpx1: acrosomal and outer dense fibre protein
- Sperm calcium ion channel; loss of motility

Table 3
Hypospermatogenesis

A. Generalized Germ Cell Loss (oligospermia)
- Bclw (antiapoptotic): progressive germ cell loss
- Aromatase (cyp19): progressive germ cell loss
- Complementation Group A: age dependent decrease
- AZFc deletions: severe oligospermia
- Hormone sensitive lipase (HSL): oligospermia
- Leydig insulin-like hormone (Ins 13): cryptorchidism
- Occludin: progressive germ cell loss
- Type 1 protein phosphatase C(2 (PP1 c(2): spermatocyte and spermatid loss.

B. Generalized Germ Cell Loss (low normal/oligospermia)
Decreased Sertoli cell numbers and “carrying capacity”
- FSH (-subunit KO
- Activin type IIA receptor KO

Table 4
Other Defects

A. Leydig Cell Agenesis, Loss or Dysfunction
- Desert Hedgehog (Dhh): Leydig cell agenesis, peritubular cell defects
- M-CSF: KO (Absent testis macrophages and absent Leydig cells
- Steroid acute regulatory protein(StAR) mutations

B. Sperm Transport Defects
- Estrogen receptor (KO: efferent duct back pressure
- Cystic fibrosis transmembrane regulator (CFTR): vas agenesis
- PEA3 (ets-transcription factor): ejaculatory dysfunction

References

1. Chillon, M., Casals, T., Mercier, B., Bassas, L., Lissens, W., Silber, S., Romey, M.C., Ruiz-Romero, J., Verlingue, C., Claustres, M. et al. (1995) Mutations in the cystic fibrosis gene in patients with congenital absence of the vas deferens. New Engl. J. Med., 332, 1475-1480.

2. Cram, D.S., O'Bryan, M.K., de Kretser, D.M. (2001) Male infertility genetics - The future. J. Androl., 22, 739-745.

3. Loveland, K.L., Schlatt, S., (1997) Stem cel factor and c-kit in the mammalian testis: lessons from Mother Nature's gene knock-outs. J. Endocrinol., 153, 337-344.

4. Giebel, L.B., Spritz R.A., (1991) Mutation of the KIT (mast/stem cell growth factor receptor) protooncogene in human piebaldism. Proc. Natl. Acad. Sci. USA, 88, 8696-8699.

5. Tut, T.G., Ghadessy, F., Trifiro, M.A., Pinsky, L., Yong, E.L. (1997) Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility. J. Clin. Endocrinol. Metab., 82, 3777-3782.

6. Dowsing, A.T., Yong, E.L., McLachlan, R.I., de Kretser, D.M., Trounson, A.O. (1999) Linkage between male infertility and trinucleotide expansion in the androgen receptor gene. Lancet, 354, 640-643.

7. Mifsud, A., Sim, C.K.S., Boettger-Tong, H., Moreira, S., Lamb, D.J., Lipshultz, L.I., Yong, E.L. (2001) Trinucleotide (CAG) repeat polymorphisms in the androgen receptor gene: molecular markers of risk for male infertility. Fertil. Steril., 75, 275-281.

8. Rajpert-De Meyts, E., Leffers, H., Petersen, J.H., Andersen, A.G., Carlsen, E., Jorgensen, N., Skakkebaek, N.E. (2002) CAG repeat length in androgen-receptor gene and reproductive variables in fertile and infertile men. Lancet, 359, 44-46.

9. Olesen, C., Hansen, C., Bendsen, E., Byskov, A.G., Schwinger, E., Lopez-Pajares, I., Jensen, P.K., Kristoffersson, U., Schubert, R., Van Assche, E. et al. (2001) Identification of human candidate genes for male infertility by digital differential display. Mol. Hum. Reprod., 7, 11-20.

10. Pennarun, G., Escudier, E., Chapelin, C., Bridoux, A.M., Cacheux, V., Roger, G., Clement, A., Goossens, M., Amselem, S., Duriez, B. (1999) Loss-of-function mutations in a human gene related to Chlamydomonas reinharditii dynein IC78 result in primary ciliary dyskinesia. Am. J. Hum. Genet., 65, 1508-1519.

11. Olbrich, H., Haffner, K., Kispert, A., Volkel, A., Volz, A., Sasmaz, G., Reinhardt, R., Hennig, S., Lehrach, H., Konietzko, N. et al. (2002) Mutations in DNAH5 cause primary ciliary dyskinesia and randomisation of left-right asymmetry. Nat. Genet., 30, 143-144.

12. Hasegawa, T., Zhao, L., Caron, K.M., Majdic, G., Suzuki, t., Shizawa, S., Sasano, H., Parker, K.L. (2000) Developmental roles of the steroidogenic acute regulatory protein (StAR) as revealed by StAR knockout mice. Mol. Endocrinol., 14, 1462-1671.

13. Print, C.G., Loveland, K., Gibson, L., Meehan, T., Stylianou, A., Wreford, N., de Kretser, D.M., Metcalf, D., Köntgen, F., Adams, J.M. et al. (1998) Apoptosis regulator Bcl-w is essential for spermatogenesis but is otherwise dispensable. Proc. Natl. Acad. Sci. USA, 95, 12423-12431.

14. Meehan, T., Loveland, K.L., de Kretser, D., Cory, S., Print, C.G. (2001) Developmental regulation of the bcl-2 family during spermatogenesis: insights into the sterility of bcl-w-/- male mice. Cell Death Differ., 8, 225-233.

15. Kumar, T.R., Wang, Y., Lu, N., Matzuk, M., (1997) Follicle stimulating hormone is required for ovarian follicle maturation but not for male fertility. Nature Genetics, 15, 201-204.

16. Wreford, N.G., Kumar, T.R., Matzuk, M.M., de Kretser, D.M. (2001) Analysis of the testicular phenotype of the follicle-stimulating hormone beta-subunit knockout and the activin type II receptor knockout mice by stereological analysis. Endocrinology, 142, 2916-2920.

What structure can tell us about funtion - sperm acrosin as a secondary zona ligand

Roy Jones PhD
Gamete Signalling Laboratory,
The Babraham Institute,
Cambridge CB2 4AT UK

INTRODUCTION

Fertilization is one of the most carefully regulated cell-cell interactions in the animal body and although many processes are involved in the formation of fully competent gametes, ultimately it is determined by the degree of compatibility between ligand and receptor molecules on the surface of sperm and egg. Identification of these molecules has proved difficult, particularly those on the sperm surface that interact with their counterparts on the zona pellucida (ZP) and eventually mediate fusion with the oolemma. In this lecture we shall outline some of the current concepts on mammalian fertilization, describe the evidence for sperm proacrosin/acrosin as a zona binding molecule and lastly, demonstrate how elucidation of the three- dimensional structure of ligand-receptor complexes can lead to a better understanding of the mechanisms involved.

CELL BIOLOGY OF FERTILIZATION IN MAMMALS

As a result of a large amount of accumulated research in several species (reviewed by Yanagimachi, 1994), fertilization can be viewed as a carefully co-ordinated series of signalling events or cross-talk between complementary gametes (Figure 1). When the fully competent spermatozoon arrives at the site of fertilization it has to penetrate through the layer of cumulous cells that surround and are attached to the zona pellucida. This is facilitated by the activity of a plasma membrane-bound form of hyaluronidase that overlies the anterior acrosomal region (steps 1-2). In most species (exceptions may be the guinea pig and rabbit) the acrosomal vesicle is still intact when the spermatozoon reaches the zona surface thereby enabling primary ligand molecules on the plasma membrane (PM) to recognize and bind to primary receptors on the ZP. In the mouse paradigm, the receptor is thought to be associated with the O-linked carbohydrate moiety of glycoprotein ZP3 (Wassarman, 1999). This binding is of sufficient affinity to tether the motile spermatozoon to the ZP surface while at the same time permitting clustering of ligand-receptor complexes that initiate Ca2+ entry and induce exocytosis of the acrosomal vesicle (steps 3-4). Point fusion of the PM with the outer acrosomal membrane leads to extensive fenestration of the acrosomal vesicle, exposes the underlying matrix and eventually the inner acrosomal membrane. In some species (e.g. hamster) an acrosomal ‘ghost’ of membranous material frequently remains attached to the ZP surface at the site of entry. At this stage another binding step takes place between the secondary receptor (ZP2) on the ZP and a secondary binding protein or ligand on the inner acrosomal membrane (step 5). Binding, however, must not be so tenacious as to prevent the sperm head from penetrating through the ZP. Partly as a result of the thrust forces generated by the motile flagellum and partly owing to an unidentified lytic-like action, the sperm head then penetrates through the ZP (step 6) and gains access to the perivitilline space. Here, a third binding step ensues that is quickly followed by membrane fusion and incorporation of the sperm head into the egg cytoplasm (step 7).

Whilst this general picture is fairly clear and orthologues of ZP receptors have been identified in several species (Rankin & Dean, 2000), the nature of the primary, secondary and tertiary ligand proteins

on spermatozoa is more contentious. A variety of candidates has been proposed and these are sufficiently disparate in their properties to raise important questions as to how they could all be necessary or integrated with one another, even allowing for redundancy within the system. Arguably, three of the most important criteria for a sperm ZP ligand molecule are (i), it should be tissue specific, (ii) it should be present in the correct place at the correct time and (iii), its mechanism of action should be compatible with the known properties of the ZP receptor. One of the few putative ligand molecules to meet these criteria is proacrosin/acrosin.

EVIDENCE FOR PROACROSIN/ACROSIN AS A SECONDARY LIGAND MOLECULE

Proacrosin is the zymogen form of the serine protease acrosin that is found within the acrosomal vesicle of all mammalian spermatozoa. It emerged as a gene duplication event from baterial trypsin approximately 1 billion years ago and is strictly tissue and organelle specific. Although its active form acrosin has long been regarded as a zona lysin to facilitate sperm penetration through the ZP, the evidence in favour of this hypothesis is less convincing than previously supposed (Bedford, 1998). Instead, it has been proposed that proacrosin/acrosin is a multifunctional protein with a major role as a secondary ligand molecule during fertilization. This supposition has arisen from experiments in which 125I-labelled pig ZP glycoproteins were used to probe western blots containing detergent-solubilised boar sperm proteins. Strong binding was obtained over a protein with a Mr of 53 kDa that was identified conclusively as proacrosin (Jones, 1990). Subsequently, substantial supportive evidence has accumulated for proacrosin/ acrosin as a secondary ZP binding molecule that can be summarised as follows.

First, ZP-proacrosin binding is non-enzymic. The mechanism of the interaction involves polysulfate groups on ZP glycoproteins forming strong ionic bonds with basic residues on the surface of the protein in a manner similar to that described for the formation of bindin-vitelline envelope complexes during fertilization in sea urchins and for binding of heparin to antithrombin III. Carbohydrates per se are not required for binding except insofar as they provide a repeating polymeric framework for the presentation of sulfate groups in the correct stereochemical alignment. This explains why some sulphated polymers (e.g. fucoidan, dextran sulfate, polyvinylsulfate) are strong competitors of ZP-proacrosin binding whereas others (e.g. heparin, chondroitin sulfates) are very weak competitors.

Second, ZP-binding activity is restricted to an internal peptide contained within residues 47 to 272 (Jansen et al., 1995). Within this peptide selected groups of basic residues on the surface of proacrosin appear to be involved in sulfate binding. These groups, which consist of various combinations of lysines, argines and histidines, show no predictable pattern in the primary sequence suggesting that it is their tertiary arrangement that is important.

Third, a polysulfonated drug called suramin not only acts as a strong competitor of ZP-proacrosin binding in solid phase binding assays, but also blocks sperm-zona binding in an ivf system (Jones et al., 1996). Suramin is a small molecular weight compound (Mr 1420) of defined chemical structure that is noteworthy for its symmetry and for the presence of 3 sulfonate groups on each of the 2 terminal naphthalene ring structures (Figure 2). Analogues that are asymmetric, or contain only 1 sulfonate group, are significantly less competitive than suramin itself suggesting that the distance or ‘span’ between the sulfonate groups is crucial for binding to proacrosin

Fourth, purified mouse ZP2 and suramin do not bind to plasma membrane-intact sperm (Howes et al., 2001). Only after permeabilisation is there significant uptake of these probes. In addition, ZP2 and suramin bind significantly less effectively to sperm from acrosin (-/-) mice than to (+/+) sperm.
Fifth, when suramin is added to mouse eggs that have been ‘pulsed’ with sperm for only 5 mins, it displaces bound sperm from the ZP surface (Howes et al., 2001), presumably because it is able to compete with ZP2 for complementary ligand molecules on sperm following the acrosome reaction. Once again, this ability is dependent on a symmetrical structure as asymmetric analogues are ineffective.

STRUCTURAL REQUIREMENTS FOR PROACROSIN-ZP BINDING

The aforegoing information strongly suggests that the 3-dimensional structure of proacrosin/

acrosin, together with the stereochemistry of projecting polysulfate groups on the carbohydrate moiety of ZP glycoproteins, are crucial for the formation of ligand-receptor complexes for retaining acrosome-reacted sperm on the ZP surface. To test these predictions, native boar and ram acrosins were crystallised and their tertiary structures determined by X-ray crystallography (Tranter et al., 2000). Of particular interest are the position of the groups of basic residues found previously to be necessary for ZP binding, namely residues His 47, Arg 50 + Arg 51 together with Arg 250, Lys 252 + Arg 253 (Jansen et al., 1998). All these residues are orientated towards the surface of the protein albeit on different sides of the active site (Figure 3). In contrast, residues Lys 73 +75 + 76 + 77, which do not seem to be obligatory for ZP binding but which still form a strong patch of positive charge, are located behind a projecting surface loop and hence are not readily available for charge-charge interactions. Interestingly, the intramolecular distance between the 2 groups of basic residues necessary for ZP binding approximates to the maximum span across the suramin molecule in an extended formation. This ability of suramin to form bridges on the surface of proteins has been demonstrated for other proteases such as elastase (Cadene et al., 1997). Suramin inhibits elastase and kallikrein strongly, trypsin weakly and chymotrypsin not at all although these enzymes belong to the same family of serine proteases and show substantial sequence homology. The difference in their ability to bind suramin probably lies in slight variations in their 3-dimensional structure. Like elastase, acrosin activity is strongly inhibited by suramin suggesting that suramin forms a bridge spanning the active site from the His 47 group to the Arg 250 group, thereby preventing access of substrate.

CONCLUSIONS

There is good evidence, therefore, to support the hypothesis that proacrosin/acrosin has a role as a secondary ligand molecule during fertilization. In this respect it fulfils all the necessary criteria mentioned earlier, the most important being that it is compatible with the known biochemical properties of the ZP receptor, which is sulphated on its carbohydrate moieties. Variations in the stereochemistry of either the sulfate groups or the position of basic residues on the surface of proacrosin could account for specificity of gamete interactions and would be subject to evolutionary selection. Other possible functions for proacrosin/acrosin, such as dispersal of the acrosomal matrix or activation of PAR-2, cannot be excluded entirely but will have to be rigorously assessed. Safe mimetics of suramin have obvious potential as antifertility agents for veterinary or human use if released near the site of fertilization, e.g. from long-acting IUDs.

REFERENCES

1. Bedford, M.J. (1998) Mammalian fertilization misread? Sperm penetration of the eutherian zona pellucida is unlikely to be a lytic event. Biol. Reprod. 59, 1275-1287.
2. Cadene, M., Duranton, J., North, A. et al. (1997) Inhibition of neutraphil serine proteases by suramin. J Biol. Chem. 272, 9950-9955.
3. Howes, E., Pascall, J.C., et al. (2001) Interactions between mouse ZP2 glycoprotein and proacrosin; a mechanism for secondary binding of sperm to the zona pellucida during fertilization. J Cell Sci. 114, 4127-4136.
4. Jansen, S., Quigley, M., Reik, W. and Jones, R. (1995) Analysis of polysulphate binding domaind in porcine proacrosin, a putative zona adhesion molecule from mammalian spermatozoa. Int. J. Dev. Biol. 39, 501-510.
5. Jansen, S., Jones, R., Jenneckens, I. et al. (1998) Site-directed mutagenesis of boar proacrosin reveals residues involved in binding of zona pellucida glycoproteins. Mol. Reprod. Dev. 51, 184-192.
6. Jones, R., Parry, R. Lo Leggio, L. et al. (1996) Inhibition of sperm-zona binding by suramin, a potential 'lead' compound for design of new antifertility agents. Hum. Reprod. 2, 5597-605
7. Jones, R. (1990) Identification and functions of mammalian sperm-egg recognition molecules during fertilization. J. Reprod. Fert. Suppl. 42, 89-105.
8. Rankin, T. and Dean, J. (2000) The zona pellucida: using molecular genetics to study the mammalian egg coat. Rev. Reprod. 5, 114-121.
9. Tranter, R., Read. J., Jones, R. et al. (2000) Effector sites in the three dimensional structure of mammalian sperm _-acrosin. Structure 8, 1179-1188.

10. Wassarman, P.M. (1999) Mammalian fertilization: molecular aspects of gamete adhesion, exocytosis and fusion. Cell 96, 175-183.
11. Yanagimachi, R. (1994) Mammalian fertilization. In: Knobil, E. and Neill, J.D. (eds). The Physiology of Reproduction. Raven Press, NY, pp 189-317.

(slide)

Figure 3. Three-dimensional structure of sperm _-acrosin with alignment of a single molecule of suramin on the same scale.

Application of the technology to patient management

Jan A.M. Kremer, MD PhD
Dept. of Ob/Gyn, University Medical Centre Nijmegen
P.O. Box 9101, 6500 HB Nijmegen, The Netherlands
j.kremer@obgyn.umcn.nl

Learning objectives
* To know what patients expect from clinicians and scientists in the future
* To know how patients should be informed about genetic aspects of male subfertility and ICSI
* To know what tests should be done before and after ICSI

Introduction
Patients, clinicians and scientists. These three interdependent groups are the main actors in the arena of medical research, and particularly in the field of genetics and male subfertility.
Scientists should know what patients expect from the exploding knowledge and technology, and subsequently focus their research-topics to these expectations. On the other hand, patients should understand what the current knowledge and technology developed by scientists mean for their personal situation.
Here lays a major task for clinicians. They should be the communicating bridge between patients and scientists. Clinicians see patients daily, listen to their problems and have an impression of their expectations. Moreover they have the medical responsibility for the potential risks of the treatment. On the other hand, clinicians can understand to a certain level what scientists are doing and can translate the current knowledge to the patients.

Firstly, I will discuss what patients expect from the scientists. What are the most interesting research-items, from the patient's point of view? Secondly, I will discuss the way clinicians should translate the current knowledge to their patients. How should we counsel them and what tests should be offered to them before and after ICSI, the most important treatment option for severe male subfertility?

Expectations of patients
The first wish of subfertile patients who visit a fertility clinic is to get pregnant. Clinicians translate this wish by offering treatments that increase the monthly chance on pregnancy.
However, getting pregnant by a safe and successful procedure is not the only wish of our patients. They also want a healthy child, preferable without future fertility problems.
Finally, most patients are interested to know the reason why they are infertile. It appeared to be frustrating if the underlying cause of the fertility problem remains unknown. They want a diagnosis.

What basic research should be done in the near future, based on these expectations and wishes?

To start with the patient's wish to know the diagnosis of their infertility, one can state that basic research on the etiology of male subfertility remains very important. More knowledge about gene-protein-structure-function, gene-expression by micro array technology, spermatogenesis genes and gene-environment interaction are necessary to give our future patients the right diagnosis.

Although ICSI has caused a revolution in the treatment of male infertility, (1) it is not the ultimate safe and successful treatment that our patients want to have. It only leads to an ongoing pregnancy in about 25% of the started cycles (2) and is often quite heavy for the woman, who gets daily injections and a transvaginal ovum pick-up. Moreover, there are risks and complications associated with this procedure, like multiple pregnancies, ovarian hyperstimulation syndrome, infection, bleeding and even death. So the first goal of research should focus on the development of more effective treatments with fewer disadvantages. In that perspective, more basic knowledge about gametes and fertilization is important. For instance better diagnostic tools on gamete level may lead to a better selection of spermatozoa for ICSI, and thus to a higher chance on pregnancy per injected oocyte. Better diagnostic tools on the level of early embryos may allow the transfer of single embryos. Than ICSI with minimal or no stimulation and transfer of single embryos become possible with fewer disadvantages and without compromising the pregnancy-rate.

Scientists also should be aware of the wish of patients for a healthy child without future fertility problems. In that perspective, again more knowledge on the (genetic) etiology of the infertility and better diagnostic tools on gamete or embryo level are important to develop. Only than, clinicians will have the tools to reduce the risks and inform their patients about the remaining problems.

Translation of the knowledge to patients
Although a lot of knowledge on this topic still has to be developed, we already know some causes of male infertility and some risks of ICSI. It is a very important task of clinicians to inform patients about the known (and unknown) risks and complications of ICSI. Clinicians should translate the knowledge developed by scientists, to their patients. Only than, the well-informed couples can make their own well-informed decision about their own reproductive future.

What should be discussed before ICSI?
First of all, the couples should be informed about all the in's and out's of ICSI itself. Secondly, clinicians should discuss the health-risks for the women (for instance OHSS, bleeding, infection), and the risks on (and consequences of) multiple pregnancies. Thirdly, clinicians should discuss intensively the partially conflicting results of the studies on congenital anomalies, chromosomal abnormalities and mental development of the children. Finally, they should inform the couple that there is a risk of transmitting the fertility problem to the offspring, particularly in the case of Y-deletions.
The best way to inform patients is probably by a combination of written and oral information.

What should be checked before ICSI?
The family history is an easy and important tool. The family history may reveal other genetic problems associated with the fertility problem. Furthermore it may give answers about heredity in cases of male infertility of unknown origin with a familiar pattern. Finally the family history may be the starting point of clinical and molecular research.
It has been known for a long time that chromosomal abnormalities are found more frequently in men with severe fertility problems. (3) Because of this higher incidence and because of the possible important consequences for the offspring, most clinicians agree that a chromosomal analysis should be performed in male ICSI candidates.
The majority of clinicians offer their patients a test for microdeletions on the Y-chromosome. Testing can be performed by PCR using a number of Sequence Tagged Sites, according to European guidelines. (4) Couples dealing with microdeletions should be counselled before they can make a decision about ICSI or no ICSI.
Mutations in the cystic fibrosis transmembrane regulator (CFTR) gene are found frequently in men with azoospermia based on congenital absence of the vas. If the woman is also a carrier, the chance to get a child with cystic fibrosis is 25%. There is no increased frequency of CFTR mutations in idiopathic male infertility, so this test only should be offered to men with a congenital absence of the vas.

What should be checked after ICSI?
Prenatal diagnosis (chromosomal analysis and advanced ultrasound examination) can be offered in ICSI-pregnancies. However, the discussion on routine prenatal chromosomal analysis after ICSI is still ongoing. The chance to detect severe chromosomal abnormalities that may lead to the decision of the couple to terminate the pregnancy is low. Advanced ultrasound may detect congenital abnormalities in mid-pregnancy. However offering this tool in all ICSI-pregnancies is disputable as well, since there are no data available that the frequency of disorders detectable by ultrasound is increased in ICSI-pregnancies. Moreover, this method may lead to unnecessary costs and anxiety.
There is a lot of dispute in the literature whether ICSI leads to a higher frequency of congenital anomalies and developmental problems (5,6). The follow-up of the children conceived after ICSI, is an important tool to give the ultimate answer to these safety-questions. Therefore every ICSI-center should have a structured follow-up procedure of the offspring.

Role for ESHRE
Ten years after the introduction of ICSI, it is time to take counseling and testing ICSI-couples seriously. This is a primary task for clinicians, who should translate the knowledge developed by scientists, to their patients. ESHRE can be helpful by making on-line, peer-reviewed, standardized information material that is understandable for patients. Moreover, ESHRE can make guidelines on counseling and testing before and after ICSI and can offer tools for a structured follow-up of ICSI-pregnancies and children.

Conclusion
Patients, clinicians and scientist are three interdependent groups, who should listen very well to each other. Clinicians have an important role as a communication bridge between patients and scientists. They should inform patients optimally about the current knowledge on genetic causes of male subfertility and risks of ICSI, and should inform scientists about the wishes and expectations of the patients.

Literature

1. Palermo G, Joris H, Devroey P, Van Steirteghem AC. Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. Lancet 1992; 340:17-8

2. Nygren KG, Andersen AN. Assisted reproductive technology in Europe, 1998. Results generated from European registers by ESHRE. European Society of Human Reproduction and Embryology. Hum Reprod. 2001; 11:2459-71

3. Tuerlings JH, de France HF, Hamers A, Hordijk R, Van Hemel JO, Hansson K et al. Chromosome studies in 1792 males prior to intra-cytoplasmic sperm injection: the Dutch experience. Eur J Hum Genet 1998; 6:194-200

4. Simoni M, Bakker E, Eurlings MC, Matthijs G, Moro E, Muller CR, Vogt PH. Laboratory guidelines for molecular diagnosis of Y-chromosomal microdeletions. Int J Androl 1999; 22:292-299

5. Bowen JR, Gibson FL, Garth LI, Saunders DM. Medical and developmental outcome at 1 year for children conceived by intracytoplasmic sperm injection. Lancet 1998; 351:1529-31

6. Bonduelle M, Joris H, Hofmans K, Liebaers I, Van Steirteghem AC. Mental development of 201 ICSI children at 2 years of age. Lancet, 1998; 351:1553

Contact info:
ESHRE Central Office
Van Akenstraat 41
1850 Grimbergen
BELGIUM
Tel: +32 (0)2 269 09 69
Fax: +32 (0)2 269 56 00
E-mail: bruno.vandeneede@eshre.com