REPRODUCTION

REVIEW

Genetics of male infertility: from research to clinic

Csilla Krausz[1,2], Antoni Riera Escamilla[1] and Chiara Chianese[1,2]

[1]*Department of Experimental and Clinical Biomedical Sciences, Centre of Excellence DeNothe, University of Florence, Viale Gaetano Pieraccini, 6, 50139, Florence, Italy and* [2]*Andrology Service, Fundació Puigvert, Instituto de Investigaciones Biomédicas Sant Pau (IIB-Sant Pau), Universitat Autònoma de Barcelona, Catalonia, Spain*

Correspondence should be addressed to C Krausz; Email: csilla.krausz@unifi.it

Abstract

Male infertility is a multifactorial complex disease with highly heterogeneous phenotypic representation and in at least 15% of cases, this condition is related to known genetic disorders, including both chromosomal and single-gene alterations. In about 40% of primary testicular failure, the etiology remains unknown and a portion of them is likely to be caused by not yet identified genetic anomalies. During the last 10 years, the search for 'hidden' genetic factors was largely unsuccessful in identifying recurrent genetic factors with potential clinical application. The armamentarium of diagnostic tests has been implemented only by the screening for Y chromosome-linked gr/gr deletion in those populations for which consistent data with risk estimate are available. On the other hand, it is clearly demonstrated by both single nucleotide polymorphisms and comparative genomic hybridization arrays, that there is a rare variant burden (especially relevant concerning deletions) in men with impaired spermatogenesis. In the era of next generation sequencing (NGS), we expect to expand our diagnostic skills, since mutations in several hundred genes can potentially lead to infertility and each of them is likely responsible for only a small fraction of cases. In this regard, system biology, which allows revealing possible gene interactions and common biological pathways, will provide an informative tool for NGS data interpretation. Although these novel approaches will certainly help in discovering 'hidden' genetic factors, a more comprehensive picture of the etiopathogenesis of idiopathic male infertility will only be achieved by a parallel investigation of the complex world of gene environmental interaction and epigenetics.

Reproduction (2015) **150** R159–R174

Introduction

Nearly 7% of men from the general population are infertile and in at least 15% of cases this condition is related to genetic disorders, including both chromosomal and single-gene alterations. Genetic causes can be detected in all major etiologic categories of male infertility (pre-testicular, testicular and post-testicular forms) and genetic tests became part of the routine diagnostic procedure in selected groups of patients (Krausz 2011). Karyotype and azospermia factor (AZF) microdeletion analyses are indicated in patients with <10 million spermatozoa/ml and <5 million spermatozoa/ml respectively (Krausz *et al*. 2014). *CFTR* gene mutation screening is performed in men affected by congenital absence of vas deferens, whereas in the case of central hypogonadism a growing number of candidate genes involved in gonadotrophin-releasing hormone receptor migration, development, secretion and response can be analyzed. After a complete diagnostic work-up (including also genetic testing), in about 40% of primary testicular failure the etiology remains unknown and is referred to as 'idiopathic infertility.' The search for 'hidden' genetic factors, especially focusing on polymorphisms, in idiopathic infertile patients were intensified in the late 1990s, since this approach turned out to be successful in some other complex multifactorial diseases (Riggs *et al*. 2014, Smith & Newton-Cheh 2015). Starting from 2009, novel approaches such as single nucleotide polymorphism (SNP) array, comparative genomic hybridization-array (array-CGH) and next generation sequencing (NGS) provided important data also on rare variants. This review is aimed at providing an overview of i) genetic risk factors including SNPs, variable number tandem repeats (VNTRs) and copy number variations (CNVs) and ii) potential causative mutations/CNVs related to idiopathic male infertility.

Genetic susceptibility factors: the candidate gene approach

Since late 1990s, the field of genetics of male infertility entered an era of intense search for genetic risk factors, mainly SNPs, VNTRs and Y chromosome-linked CNVs. The results obtained up to 2007 have been summarized in the meta-analysis by Tüttelmann *et al*. (2007), who reported significant association with impaired

ISSN 1470–1626 (paper) 1741–7899 (online)

DOI: 10.1530/REP-15-0261

Online version via www.reproduction-online.org

spermatogenesis only for two genetic factors: a partial AZFc deletion (gr/gr deletion) and the rs1801133 (c.677C>T) variant in the *MTHFR* gene. At that time, for many other SNPs, either only single studies were available or results from different laboratories were discordant (Nuti & Krausz 2008).

We herein review the existing literature via a search in the PubMed database of case–control studies published since 2008. The following keywords were used to select eligible studies: 'genetic risk factor (s)' AND 'male infertility.' Additionally, all identified gene/polymorphism combinations were searched individually (e.g., 'FASLG' and 'male (in)fertility'). Data were extracted from single papers and are summarized in Tables 1, 2 and 3 and Supplementary Table 1.

As in other fields of medicine, targeted search for SNPs or gene mutations is based on the candidate gene approach. This approach has been facilitated by an increasing body of information from model organisms, expression analyses (transcriptomic and proteomic) in relationship with spermatogenesis and, together with data produced by Genome-Wide Association Studies (GWAS) (Tables 2 and 3), represents the major source for genetic studies in humans. A minority of SNPs ($n=28$) studied before 2008 have been the objects of subsequent publications, whereas the large majority, listed in Table 1, are new entries ($n=286$). A total of 314 SNPs have been reported in 123 genes. Approximately 70% of SNPs are related to genes with common cell function but with predicted relevance in germ cells, such as apoptotic process, DNA repair, detoxification of environmental molecules, response to reactive oxygen species and so on. Indeed, the best candidate genes are those with specific expression in germ cells or those that have specific spermatogenic function or play important roles in meiosis or endocrine regulation of the testis (Table 1). Data in existing literature are rarely concordant, and for many SNPs ($n=269$), only single studies are available. To date, meta-analyses are available for ten genes: *AR, CYP1A1, DAZL, ESR1, ESR2, MTHFR, NOS3, POLG, TP53* and *USP26*. Although data remains largely controversial, ethnic/geographic origin seems to play an important role in the phenotypic expression of polymorphisms in the *MTHFR, ESR1/ESR2*, *NOS3* and *DAZL*. Data remains inconclusive for *CYP1A1* and *AR* genes, whereas a lack of association with male infertility has been clearly demonstrated for polymorphisms related to *TP53, USP26* and *POLG*. Although reliability of the presently available meta-analyses is largely limited by the heterogeneous inclusion criteria used for patients and controls selection, in this review we attempt to provide a short description of those SNPs that according to the latest meta-analyses result significantly associated with spermatogenic failure.

Tüttelmann *et al.* (2007) reported that the c.677C>T variant in the *MTHFR* (methylenetetrahydrofolate reductase (*NAD(P)H*) gene was the only one showing significant association with male infertility. The *MTHFR* gene is located on chromosome 1p.36.22, encodes an enzyme that produces 5-methyltetrahydrofolate and is in involved in folate metabolism. Folate is necessary for the preservation of genome integrity due to its role in DNA synthesis, repair and methylation, and it has been predicted that its deficiency may lead also to male infertility. The c.677C>T variant impairs the enzyme activity by 35% in heterozygosis and by 70% in homozygosis (Frosst *et al.* 1995). The conclusion presented by Tüttelmann *et al.* (2007) stimulated further studies, which led to controversial results and to novel meta-analyses (Gupta *et al.* 2011, Wei *et al.* 2012, Wu *et al.* 2012, Weiner *et al.* 2014, Gong *et al.* 2015). Interestingly, there is discordance even between the five meta-analyses, with some reporting an association (Tüttelmann *et al.* 2007, Gupta *et al.* 2011, Wu *et al.* 2012) and others reporting a lack of association (Wei *et al.* 2012, Weiner *et al.* 2014). The last meta-analysis (Gong *et al.* 2015), which included 26 published studies (5575 cases and 5447 controls from Asian, African and Caucasian populations), indicated that the *MTFHR* variant is associated with AZ (AZ) (OR=1.36, 95% CI: 1.18–1.55, $P=0.000$) and oligoasthenoteratozoospermia (OAT) (OR=1.35, 95% CI: 1.11–1.64, $P=0.003$), but not with oligozoospermia. Finally, a second SNP in the *MTHFR* gene has also been the object of numerous studies but with similar discordant results. Rs1801131, also known as 1298C>A, is a missense polymorphism found in exon 7 that also reduces *MTHFR* activity, though apparently less severely than C677T (Van der Put *et al.* 1998). The meta-analysis of seven studies with a total of 1633 cases and 1735 controls from different ethnic groups shows that the polymorphism is significantly associated with azoospermia (OR=1.12, 95% CI=1.00–1.26) but not with OAT (Shen *et al.* 2012).

Overall, for both SNPs the conferred susceptibility to AZ and OAT is modest, implying a marginal biological role for this SNP in infertility. Controversies might depend on different ethnic origin (variant frequency does differ among different populations), and the penetrance of this mutation is likely to be affected by diet, e.g., subjects carrying the variant may have a major risk for male infertility in cases of low folate intake. Consequently, it could be of interest to test for these SNPs in relationship to the responsiveness to folate supplementation, i.e., to select potential 'responders' through a pharmacogenetic approach.

Other SNPs that have been objects of investigation occur in the estrogen receptor 1 (*ESR1*) and estrogen receptor 2 (*ESR2*) genes. Estrogens are predicted to play an important role in the male reproductive tract, and both the deficit and the excess of estrogens can alter sperm production and maturation (Atanassova *et al.* 1999, Hess 2003). Three different receptor isoforms ERα, and ERγ are known. The *ESR1* gene on 6q25 codifies for ERα, a 595 amino acid receptor. The *ESR2* gene is

Table 1 Summary of case-control studies focusing on gene polymorphisms since 2008. SNPs related to genes with (A) common cell function, (B) specific spermatogenic function, (C) endocrine function. Further details are given in the Supplementary Table 1, see section on supplementary data given at the end of this article.

Gene name	Cases + controls	Country of origin	Association
(A) Common cell function			
ABCB1[a]	162 + 191	Poland	YES
ABLIM1[a]	3608 + 5909	China	YES
AHR	991 + 1256	China; Estonia; Iran; Japan	YES**
AHRR	235 + 324	Estonia; Japan	DISCORDANT
APOB	604 + 501	Slovenia; India	DISCORDANT
ARNTL[a]	589 + 444	Slovenia, Serbia	NO
ATM	809 + 816	China	DISCORDANT
BCL2[a]	1653 + 2329	China	YES
BHMT[a]	153 + 184	Sweden	NO
BRCA2	820 + 830	China	YES**
CAT	885 + 839	China; France; Iran	DISCORDANT
CDC42BPA[a]	3608 + 5909	China	YES
CHD2[a]	1653 + 2329	China	NO
CLOCK[a]	517 + 444	Slovenia	YES
CRISP2[a]	92 + 176	Australia	NO
CYP1A1	1060 + 1225	Meta-analysis	YES
CYP17A1[a]	456 + 465	Korea	YES
CYP26B1[a]	719 + 383	China	NO
EPSTI1	917 + 2015	Japan	DISCORDANT
ERCC1[a]	202 + 187	China	NO
ERCC2	202 + 187	China	NO
ETV5[a]	204 + 296	Australia, USA	YES
FAS	547 + 571	China; India; Turkey	NO
FASLG	447 + 532	Albania, Macedonia; China; Turkey	NO
FOLH1[a]	153 + 184	Sweden	NO
GNAO1[a]	1653 + 2329	China	YES
GPX1	690 + 649	China; France	NO
HLA-DRA	4508 + 7588	China; Japan	YES
JMJDIA[a]	136 + 161	Albania, Macedonia	NO
KLK2[a]	218 + 220	Korea	YES
LIG4[a]	580 + 580	China	YES
LOC203413	623 + 530	Albania, Macedonia; Japan	NO
LRWD1	130 + 100	Japan	NO
MAS1L/UBD	917 + 2015	Japan	NO
MCT2 (SLC16A7)[a]	471 + 265	Korea	YES
MDM2[a]	580 + 580	China	YES
MLH1[a]	1292 + 480	China	NO
MLH3	1454 + 640	China	YES**
MSH4[a]	1292 + 480	China	NO
MSH5	1454 + 640	China	YES
MTHFD1	428 + 533	Sweden; Russia	NO
MTHFR	5575 + 5447	Meta-analysis	YES
MTR	713 + 739	Brazil; China; Poland	NO
MTRR	1790 + 1622	Brazil; China; France; Jordania; Korea; Poland; Sweden	DISCORDANT
NFE2L2 (NRF2)[a]	336 + 295	China	YES
NOS1[a]	580 + 580	China	NO
NOS2[a]	580 + 580	China	NO
NOS3	2019 + 1509	Meta-analysis	DISCORDANT
NQO1[a]	580 + 580	China	NO
OR2W3	623 + 530	Albania, Macedonia; Japan	DISCORDANT
PACRG[a]	610 + 156	Australia	YES
PARP1[a]	317 + 231	China	YES
PCFT1[a]	153 + 184	Sweden	NO
PEMT[a]	153 + 184	Sweden	YES
PEX10	2369 + 2946	China; Japan	NO
PMS2[a]	1292 + 480	China	YES
POLG	2463 + 1480	Meta-analysis	NO
PON1	1037 + 1094	China; Greece; Iran; Slovenia	DISCORDANT
PON2	270 + 320	Greece; Iran	DISCORDANT
PSAT1	917 + 2015	Japan	DISCORDANT
RAG1[a]	580 + 580	China	YES
RFC1[a]	153 + 184	Sweden	NO
RGS9[a]	3608 + 5909	China	NO
SHMT1	153 + 184	Sweden	NO
SFRS1[a]	962 + 1931	China	NO
SFRS2[a]	962 + 1931	China	NO
SFRS3[a]	962 + 1931	China	NO
SFRS4[a]	962 + 1931	China	NO
SFRS5[a]	962 + 1931	China	NO
SFRS6[a]	962 + 1931	China	YES
SFRS7[a]	962 + 1931	China	NO
SFRS9[a]	962 + 1931	China	NO
SIRPA	1402 + 1172	China	YES**
SIRPA-SIRPG[a]	490 + 1167	China	NO
SIRPG	1402 + 1172	China	DISCORDANT
SOD2	690 + 649	China; France	DISCORDANT
SOD3[a]	580 + 580	China	NO
SOX5	2987 + 3526	China; Japan	DISCORDANT
TAS2R38	623 + 530	Macedonia, Albania and Japan	NO
TCblR[a]	153 + 184	Sweden	YES
TCN2[a]	153 + 184	Sweden	NO
TMEM132E[a]	3608 + 5909	China	NO
TNF[a]	780 + 260	India	YES
TP53	1134 + 1545	Meta-analysis	NO
UBR2[a]	30 + 80	Japan	YES
USP26	1716 + 2597	Meta-analysis	NO
USP8	917 + 2015	Japan	DISCORDANT
XPC[a]	252 + 288	China	NO
XRCC2[a]	580 + 580	China	NO
XRCC3[a]	580 + 580	China	NO
XRCC4[a]	580 + 580	China	NO
XRCC5[a]	580 + 580	China	NO
(B) Specific spermatogenic function			
BRDT	259 + 343	Albania, Macedonia; Israel	NO
DAZL	2715 + 1835	Meta-analysis	DISCORDANT
EPPIN[a]	473 + 198	China	YES
H2BFWT	851 + 445	China; Korea	YES
HORMAD1	391 + 448	China; Japan	YES**
HORMAD2[a]	361 + 368	China	NO
MOV10L1[a]	30 + 70	Iran	NO
NANOS1[a]	719 + 383	China	NO
PIWIL1[a]	490 + 468	China	NO
PIWIL2[a]	490 + 468	China	NO
PIWIL3[a]	490 + 468	China	NO
PIWIL4[a]	490 + 468	China	NO
PRDM9[a]	309 + 377	China	NO
PRM1	851 + 955	China; Iran; Japan; Spain	YES**
PRM2	525 + 648	China; Japan	NO
PRMT6	2369 + 2946	China; Japan	NO
REC8[a]	96 + 96	USA	NO

Table 1 *Continued.*

Gene name	Cases+ controls	Country of origin	Association
SEPT12	290+480	Japan; Taiwan	DISCORDANT
SPATA17[a]	38+96	Japan	YES
SPO11	186+167	China; Iran	DISCORDANT
STRA8[a]	719+383	China	YES
TEX15	445+538	Albania, Macedonia; China	NO
TSSK4[a]	372+220	China	NO
TSSK6[a]	519+359	China	NO
UBE2B	568+612	China and India	YES*[a]
YBX2[a]	326+210	China	YES
(C) Endocrine function			
AR	2084+1831	Meta-analysis	YES
ESR1	1576+1777	Meta-analysis	DISCORDANT
ESR2	2815+3178	Meta-analysis	DISCORDANT
INSR	624+530	Albania, Macedonia; Japan	NO
MSMB[a]	338+382	China	YES
SRD5A2[a]	132+111	Estonia	NO

Underlined, gene polymorphisms evaluated in meta-analyses comprising study populations with different ethnic/geographic origins and association description refers to the global meta-analysis results; YES, SNP is associated in all studies; YES**, multiple SNPs studied in the gene by different authors, but specific SNPs analyzed in a single study result as associated to male infertility; DISCORDANT, the same SNP analyzed in different studies show discordant results; NO, SNP shows no association in any study.
[a]Gene analyzed by a single study. Alternative gene names appearing in other studies are reported in brackets.

located on chromosome 14q23-24 and codifies for ERβ, a protein with 530 amino acids. Both receptors are highly expressed in human testicular germ cells. Regarding *ESR1,* the two most studied SNPs are rs2234693 (also known as PvuII) and rs9340799 (known as XbaI), both located in intron 1 (c.453-397T>C and c.453-351A>G respectively). Although a relationship between these SNPs and *ESRs* gene/protein function and stability has been proposed, their exact effect remains unclear. The last meta-analysis performed so far involves 12 studies comprising from 736 to 1418 infertile cases and 841–1601 controls depending on the type of analyzed SNP (Ge *et al*. 2014). The meta-analysis includes azoospermic, oligozoospermic and oligoasthenozoospermic (OAZ) and OAT patients of different ethnic and geographic origin. According to this analysis, ethnic background plays an important role in the biological effect of the variants. For instance, the minor allele C of rs2234693 (c.453-397T>C) seems to show a protective effect in the Asian population (C allele vs T allele OR=0.78, 95% CI: 0.64–0.96; CC vs TT, OR=0.61, 95% CI: 0.40–0.93), whereas in Caucasians it is associated with an increased risk for infertility (CC vs CT+TT: OR=1.52, 95% CI: 1.05–2.22). As far as the XbaI SNP (c.453-351A>G), the G allele is associated with a decreased risk, according to the dominant model in the Asian population, whereas no association was found in Caucasians. A similar situation was encountered also for the SNP rs1256049 in *ESR2* (c.984G>A), which according to the recessive model is associated with a decreased risk in Asian populations, whereas in Caucasian men it is associated with an increased risk for male infertility according to the dominant model. Finally, rs4986938 (c.1406+1872G>A) mapped on *ESR2* does not affect male fertility in any population. These results show again the importance of the patients' ethnic origin and their genetic background in modulating the effect of a given variant. Controversies may also derive from the different level of exposition to endocrine disrupters, which also interact with these receptors and alter testis development and function. It is therefore plausible that a more pronounced effect of these SNPs can be observed only in relationship with a high level of exposure to these environmental factors.

As for the nitric oxide synthase 3 (*NOS3* or *eNOS*) gene, three principal SNPs have been studied in relationship with male infertility: rs1799983 (c.894T>G in the exon 8), rs2070744 (c.-786C>T in the promoter region) and rs61722009 (27 bp VNTR polymorphisms in the intron 4, also known as 4a4b polymorphisms). *NOS3* is located on chromosome 7q36.1 and produces nitric oxide (NO), which is implicated in several cellular functions such as vascular smooth muscle relaxation through a cGMP-mediated signal transduction pathway, but also predicted to have an important role in fertility, including sperm motility and maturation, as well as germ cell apoptosis in the testis (Zini *et al*. 1996, Lee & Cheng 2008). The *eNOS* rs2070744 variant is associated with reduced promoter activity, suppressed eNOS transcription and decreased NO generation (Dosenko *et al*. 2006). There is also a trend for diminished eNOS enzyme activity in *eNOS* rs1799983 SNP carriers (Wang &.Mahaney 1997). The VNTR within intron 4 of the *eNOS* gene accounts for >25% of basal plasma NO generation, suggesting that this gene might have an important role in NO-mediated physiology (Wang *et al*. 1997). The first case–control study related to fertility analyzed the three SNPs in a cohort of 371 patients and association was found only between the 4a4b variant and sperm morphology (Yun *et al*. 2008). Subsequently, relatively small studies from Italy, China, Iran and Brazil reached discordant results (Buldreghini *et al*. 2010, Safarinejad & Shafiei 2010, Bianco *et al*. 2013, Yan *et al*. 2014). Finally, Song *et al*. (2015) performed a meta-analysis on 2018 infertile patients (from eight studies, including their own) and concluded that only c.-786C>T and 4a4b were significantly associated with male infertility in both the Asian and Caucasian populations (OR=1.53, 95% CI=1.10–2.22 and OR=3.24, 95% CI=2.49–4.22 respectively). Indeed, these SNPs are promising and merit further investigations in order to define their potential clinical relevance.

The deleted in azoospermia-like (*DAZL*) gene is an autosomal homologue of the Y-chromosomal *DAZ* (deleted in azoospermia) gene cluster and maps to chromosome 3p24 (Yen *et al*. 1996). As the other family

Table 2 Summary of GWAS results. SNPs and related genes described as significantly associated in GWA Studies.

Aston & Carrell (2009)		Aston *et al.* (2010)[a]		Hu *et al.* (2012)		Zhao *et al.* (2012)		Kosova *et al.* (2012)	
SNP associated	Gene related	SNP associated	Gene related	SNP associated	Gene related	SNP associated	Gene related	SNP associated	Gene related
rs1399645	*NXPH2*	rs763110	*FASLG*	rs12097821	*PRMT6*	rs3129878	*HLA-DRA*	rs10966811	*TUSC1*
rs2063802	*NXPH2*	rs5911500	*LOC203413*	rs2477686	*PEX10*	rs498422	*C6orf10/BTNL2*	rs7867029	*PSAT1*
rs4954657	*NXPH2*	rs10246939	*TAS2R8*	rs10842262	*SOX5*			rs12870438	*EPSTI1*
rs11707608	*CNTN3*	rs3088232	*BRDT*					rs7174015	*USP8*
rs2976084	*CNTN3*	rs323344	*TEX15*					rs10129954	*DPF3*
rs3105782	*MASP1*	rs323345	*TEX15*					rs680730	*DSCAML1*
rs4484160	*PROK2*	rs5764698	*SMC1B*					rs11236909	*TSKU/LRRC32*
rs9814870	*ARL6*	rs1801131	*MTHFR*					rs10488786	*ARHGAP42*
rs9825719	*NSUN3*	rs631357	*KIF17*					rs724078	*MAS1L/UBD*
rs2290870	*ATP8A1*	rs35397110	*USP26*						
rs4343755	*GNPDA2*	rs34605051	*JMJD1A*						
rs4695097	*GNPDA2*	rs2030259	*JMJD1A*						
rs4541736	*LRFN2*	rs11204546	*OR2W3*						
rs1545125	*COBL*	rs2059807	*INSR*						
rs215702	*LSM5*								
rs6476866	*SLC1A1*								
rs10841496	*PDE3A*								
rs10848911	*EFCAB4B*								
rs12920268	*MAF*								
rs2032278	*GALR1*								
rs608020	*SALL4*								

[a]Aston *et al.* (2010) analyzed a total of 172 SNPs including also 84 SNPs from Aston & Carrell (2009).

members (*DAZ* and *BOLL*), this gene encodes RNA binding proteins with important roles in spermatogenesis (Yen 2004). One of the most studied SNPs is rs121918346, a missense variant that changes threonine 54 to an alanine on exon 3. The last meta-analysis comprised 13 studies with a total of 2715 cases and 1835 controls from different ethnic origins and concluded that the variant was significantly associated with male infertility exclusively in Chinese men (Chen *et al.* 2015). This finding is in line with the conclusion of the first Caucasian study that considered this polymorphism as 'an example of remarkable ethnic differences' for its effect on predisposing carriers to spermatogenic failure (Becherini *et al.* 2004).

The androgen receptor *(AR)* gene also contains two polymorphic sites in the N-terminal trans-activation domain of the receptor: a polyglutamine tract – $(CAG)_n$ – and a polyglycine tract – $(GGC)_{n,}$, which were objects of many publications related to male infertility (for review see Davis-Dao *et al.* (2007) and Nenonen *et al.* (2011)) The $(CAG)_n$ length normally ranges between six and 39 repeats in the general population, with a median value that varies according to the ethnicity (21–22 in White Caucasian, 19–20 in African–American, 22–23 in Asian, 23 in Hispanic populations). The originally described inverse relationship between CAG repeat length and the receptor trans-activation led to the hypothesis that longer CAG repeat conferred a higher risk for a series of androgen-dependent diseases, including infertility and cryptorchidism (Tut *et al.* 1997). The first meta-analysis based on 33 publications in 2007 was unable to find a cut-off value above which infertility risk is increased (Davis-Dao *et al.* 2007). A more recent meta-analysis has proposed an alternative way of analysis based on the 'optimal range' hypothesis, which derives from novel functional studies reporting that the AR activity was actually higher in the presence of a determined number of CAG (Nenonen *et al.* 2011). Therefore, according to this hypothesis either a longer or a shorter CAG tract might have a negative effect on the receptor function. Although Nenonen *et al.* (2011) were able to demonstrate a significant association between the length of this polymorphism below or above the 'optimal range' and impaired sperm production (CAG <22: $P=0.03$, OR $=1.18$ 95% CI: 1.02–1.39; for CAG >23: $P=0.02$, OR $=1.22$, 95% CI 1.03–1.44), the role of CAG repeats in male infertility is probably more complex than it has been previously considered. More functional and clinical studies are needed before the introduction of this polymorphism into the diagnostic setting.

The CYP1A1 (cytochrome P450, family 1, subfamily A, polypeptide 1) is located on chromosome 15q24.1 and encodes a member of the cytochrome P450 superfamily. The cytochrome P450 proteins are mono-oxygenases that catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids. *CYP1A1* encodes a 522-aminoacide protein that, among its functions, is involved in the metabolism of polycyclic aromatic hydrocarbons into their biologically active intermediates that have potential reproductive toxicity in men (McManus *et al.* 1990). The rs4646903 variant, a T>C substitution in 3'UTR of

Table 3 Summary of GWAS replication studies for SNPs and related genes (including SNPs presenting significant or borderline association in the original GWAS).

Reference	SNPs analyzed	Gene related
Follow-up Aston *et al.* (2010)		
Plaseski *et al.* (2012)[a]	rs5911500[b]	*LOC203413*
	rs11204546[b]	*OR2W3*
	rs3088232[b]	*BRDT*
	rs2059807	*INSR*
	rs10246939	*TAS2R8*
	rs34605051	*JMJD1A*
	rs323344	*TEX15*
	rs323345	*TEX15*
	rs763110	*FASLG*
Chihara *et al.* (2015)	rs11204546[b]	*OR2W3*
	rs5911500	*LOC203413*
	rs10246939	*TAS2R8*
	rs2059807	*INSR*
Follow-up Hu *et al.* (2012)		
Xu *et al.* (2013)	rs3197744[b]	*SIRPA*
	rs11046992[b]	*SOX5*
	rs146039840	*SOX5*
	rs1129332	*PEX10*
	rs3791185	*PRMT6*
	rs2232015	*PRMT6*
	rs1048055	*SIRPG*
Lu *et al.* (2014)	*rs1048055*[b]	*SIRPG*
	rs2281807	*SIRPG*
	rs11046992	*SOX5*
	rs146039840	*SOX5*
Zou *et al.* (2014)	rs10842262[b]	*SOX5*
	rs12097821	*PRMT6*
	rs2477686	*PEX10*
Hu *et al.* (2014)[c]	rs7194[b]	*HLA-DRA*
	rs7099208[b]	*ABLIM1*
	rs13206743[b]	*MIR133BL17A*
	rs3000811[b]	*CDC42BPA*
Sato *et al.* (2013)	rs12097821	*PRMT6*
	rs2477686	*PEX10*
	rs10842262	*SOX5*
	rs6080550	*SIRPA-SIRPG*
Follow-up Hu *et al.* (2012), Zhao *et al.* (2012)		
Tu *et al.* (2014)	rs3129878[b]	*HLA-DRA*
	rs12097821	*PRMT6*
	rs10842262	*SOX5*
	rs2477686	*PEX10*
Follow-up Zhao *et al.* (2012)		
Jinam *et al.* (2013)	rs3129878[b]	*HLA-DRA*
	rs498422	*C6orf10/BTNL2*
Follow-up Kosova *et al.* (2012)		
Sato *et al.* (2015)	rs7867029[b]	*PSAT1*
	rs7174015[b]	*USP8*
	rs12870438[b]	*EPSTI1*
	rs724078	*MAS1L/UBD*

[a]SNPs in this study are not significantly associated after Bonferroni correction. [b]SNPs described as significantly associated. [c]Only SNPs described as significantly associated to male infertility are listed (in the study, a total of 77 SNPs originated from the Hu *et al.* (2012) paper were screened).

CYP1A1 gene has been associated with increased transcript half-life and therefore increased enzyme activity resulting in elevated levels of activated metabolites (Manfredi *et al.* 2007). This SNP has been associated with different types of cancers (Salnikova *et al.* 2013, Abbas *et al.* 2014), further supporting their biological importance. Studies focusing on the role of this SNP in male infertility overall produced discordant results even in the same ethnic groups. Despite discrepancies, the last meta-analysis performed on a total of 1060 cases and 1225 controls concluded for a significant association between the variant and male infertility reaching the highest risk's entity according to the homozygous model (OR = 2.18, 95% CI: 1.15–4.12) (Luo *et al.* 2014). However, since only two out of six studies report it as a significant susceptibility factor, this meta-analysis awaits further confirmation. Given the biological function of this gene, differences in exposure to environmental factors may also influence the outcome of single studies; lack of information about careful matching of important variables such as drug and alcohol intake and life-style factors between patients and controls may well be responsible for controversies.

Apart from the meta-analyses focusing on the ten genes, in case of multiple studies analyzing the same SNPs/gene, results are almost constantly controversial and even if association is found generically with 'infertility,' the subgroup analysis shows differences (Supplementary Table 1). An example is the rs7885967 (c.-9C > T) of the *H2BFWT* (H2B histone family, member W, testis-specific) gene encoding for a testis-specific histone with an essential role during meiotic chromatin reorganization (Gineitis *et al.* 2001). This SNP maps to the 5′UTR of *H2BFWT* and has been demonstrated to affect the translation of the protein (Lee *et al.* 2009). The two case–control studies found significant association (with moderate OR ranging from 1.51–1.88) with completely different semen phenotypes: azoospermia in the Chinese population (Ying & Scott 2012) whereas lack of association with azoospermia and association with non-azoospermia (a heterogenous group of oligo/astheno/teratozoospermic men) in the Korean study (Lee *et al.* 2009). Such contradictory results clearly discourage further studies on this SNP.

The unique example of a polymorphism with fully concordant results in more than one relatively large independent study populations is related to the *MSH5* gene (rs2075789). The mutS homolog 5 (*MSH5*) encodes a member of the mutS family of proteins that are involved in DNA mismatch repair and apoptosis. *Msh5* knockout mice present sterility due to the defect in resolving meiotic chromosomal crossovers (Edelmann *et al.* 1999) Yeast two-hybrid analysis demonstrated that the SNP rs2075789 impairs interaction between MSH4 and MSH5 proposing a functional effect (Yi *et al.* 2005). The two independent studies that include a total of 1454 cases and 640 controls from the Chinese population report a similar risk's entity for homo/heterozygous minor allele carriers compared to WT homozygous carriers (OR = 2.51; 95% CI = 1.43–4.40 and OR = 1.83, 95% CI = 1.32–2.55, by Xu *et al.* (2010) and Ji *et al.* (2012) respectively). Although this is a promising candidate SNP, its importance remains limited until new data are available in other populations.

Genetic susceptibility factors: GWAS and SNPs

All the genetic risk factors discussed above originate from the candidate gene approach, which is based on the analysis of genes/polymorphisms with predicted or known function in spermatogenesis. Given the relatively poor outcome of these studies, much expectation was given to whole genome analysis. Gene discoveries from GWAS have been successful for several diseases and helped unravel pathways important for a certain biological process (Visscher *et al*. 2012) Overall, four GWAS based on SNP-arrays are available in the literature and are summarized in Table 2 (Aston & Carrell 2009, Hu *et al*. 2012, Kosova *et al*. 2012, Zhao *et al*. 2012). The first study by Aston and Carrell (2009) analyzed 370 000 SNPs in 92 oligozoospermic and non-obstructive azoospermic (NOA) patients and 80 healthy controls and found 21 SNPs associated with azoospermia or oligozoospermia. Due to the prohibitively high cost of the array studies in 2009, the study population size was clearly underpowered and the associations reported did not reach genome-wide significance. This pioneer work was followed by two large, properly powered Chinese GWAS, which reported a number of SNPs with stringent *P* value $<1\times10^{-8}$. Hu *et al*. (2012) analyzed 2927 individuals with NOA and 5734 controls from Han Chinese population and found a few SNPs predisposing to NOA in *PRMT6, PEX10* and *SOX5* genes. The second study analyzed 2226 NOA patients and 4576 controls in the same population and reported significant associations with SNPs mapping to two regions: *HLA-DRA* and *C6orf10/BTNL2* (Zhao *et al*. 2012). Despite meeting requirements for genome-wide significant results, no overlapping SNPs were observed between these two large studies. Finally, in the same year Kosova *et al*. (2012) analyzed 269 Hutterite men and 123 men from Chicago with diverse ethnic background, and described nine SNPs associated with reduced fertility or impaired sperm parameters, but in this case also no SNPs overlapping with the previous three GWAS were reported (Table 2).

Subsequently, SNPs reported as significantly associated or with borderline *P* values in the above GWAS were analyzed in independent study populations with variable success (Table 3). Findings on the majority of candidate SNPs were not confirmed by the replication studies, and the few SNPs that show association either confer a moderate risk for impaired sperm production or loose significance after Bonferroni correction (for instance, *OR2W3, BRDT*). Interestingly, the SNP reported in *SIRPA/G* (rs6080550) with borderline significance in one of the GWAS (Hu *et al*. 2012) was not confirmed in the follow-up studies, but following re-sequencing of the *SIRPA* gene, another SNP (rs3197744) was identified as a significant susceptibility factor for oligozoospermia with OR=4.62 (95% CI= 1.58–13.4 *P*=0.005) (Xu *et al*. 2013) Similarly, the re-sequencing of *SIRPG* also provided an interesting candidate SNP (rs1048055) with similarly high OR for NOA (OR=3.93, 95% CI=1.59–9.70 $P=3.00\times10^{-3}$) (Lu *et al*. 2014). Both genes are members of the signal-regulatory-protein (SIRP) family and belong to the immunoglobulin superfamily, and when they bind to CD47 can induce cell apoptosis (Brooke *et al*. 2004). According to the above data, *SIRPA/G* can be considered as promising candidate genes for spermatogenic impairment and furtherer investigations.

The *HLA-DRA* gene-related SNPs turned out to be the most promising, since highly significant association with NOA was found in the GWAS of Zhao *et al*. (2012) and in four case–control studies in Chinese and Japanese populations (Tsujimura *et al*. 2002, Jinam *et al*. 2013, Hu *et al*. 2014, Tu *et al*. 2014). *HLA-DRA* gene is a member of class II genes and encodes the alpha chain of HLA-DR and heterodimerizes with β chains (HLA-DRBs) and plays an important role in the immune system by presenting peptides on the cell surface of antigen-presenting cells. Three variants have been described with significant association with male infertility in Japanese and Chinese populations (Zhao *et al*. 2012, Jinam *et al*. 2013, Hu *et al*. 2014, Tu *et al*. 2014): rs3129878, rs7194 and rs7192. The variant rs7194 is in linkage disequilibrium with rs7192 and is located on 3′UTR. It was predicted to map to the has-miR-6507-3p binding site and may play an important role during transcription by influencing *HLA-DRA* expression level through microRNA-mediated post-transcriptional regulation (Lin *et al*. 2015). As for rs7192, it is a missense variant (L242V) located in exon 4, which encodes part of the DRA α-chain cytoplasmic domain (Neefjes *et al*. 2011). This SNP might alter interactions with β-chain or ubiquitin E3 ligases, which control the cell-surface expression of class II MHC proteins (Gueant *et al*. 2015). Finally, rs3129878 maps to intron 1 and its putative effect is not yet clarified. These polymorphisms have been already described as susceptibility factors for a number of autoimmune diseases, therefore it has been hypothesized that they might mediate the response to testicular micro-environmental antigens and therefore may elicit autoimmune inflammatory responses leading to azoospermia (Hu *et al*. 2012). It would be interesting to study this polymorphism also in Caucasians and in subgroups of patients with previous history of urogenital inflammation, especially orchiepididymitis.

Rare variants: gene re-sequencing studies

Besides the polymorphisms described above, many re-sequencing studies of candidate spermatogenesis genes have been also published. Although many genes are known to be essential for gametogenesis, there are surprisingly few monogenic mutations that have been conclusively demonstrated to cause human spermatogenic failure. The majority of mutations identified are in

heterozygosis and therefore the demonstration of a cause-effect relationship remains difficult. In addition, functional studies are lacking in a large majority of the cases. Some of the most promising mutations, for which also functional studies were performed, have been identified in the following genes: i) *HSF2* (Mou *et al*. 2013) and *SOHLH1* (Choi *et al*. 2010) reported in NOA men; ii) *NANOS1* (Kusz-Zamelczyk *et al*. 2013) and *NR5A1* (Bashamboo *et al*. 2010) reported in NOA and oligozoospermic patients; iii) Yatsenko *et al*. 2006), *GALNTL5* (Takasaki *et al*. 2014) and *SEPT12* (Kuo *et al*. 2012) identified in oligo or OAT men. All the above genes are autosomal and the reported mutations are in heterozygosis. Whether these mutations are fully responsible for the given phenotypes (dominant effect) or are acting in synergy with other yet unidentified heterozygous mutations in genes with similar function (oligogenic model) remains to be defined.

Thanks to the diffusion of NGS platforms, testing for a large panel of candidate genes in large group of patients and controls has now became an affordable approach. The first NGS-based, candidate gene panel study has been recently performed in a Chinese case–control setting including 757 NOA patients and 709 fertile males (Li *et al*. 2015), Using the HiSDefault 2000 platform, they sequenced a total of 650 infertility-related genes and described a significant excess of rare, non-silent variants in genes that are key epigenetic regulators during spermatogenesis such as *BRWD1, DNMT1, DNMT3B, RNf17, UBR2, USP1* and *USP26.* The authors do not provide detailed information about the exact genotype of the variants, but apparently 'most of the non-silent variants in these genes in the sporadic NOA patients were heterozygous.' As *USP26* is located on the X chromosome, the reported variants are hemizygous. Given that these genes are involved in similar biological function, the hypothesis about a synergic action of heterozygous mutations is plausible. However, functional analyses are still needed in order to support this hypothesis,

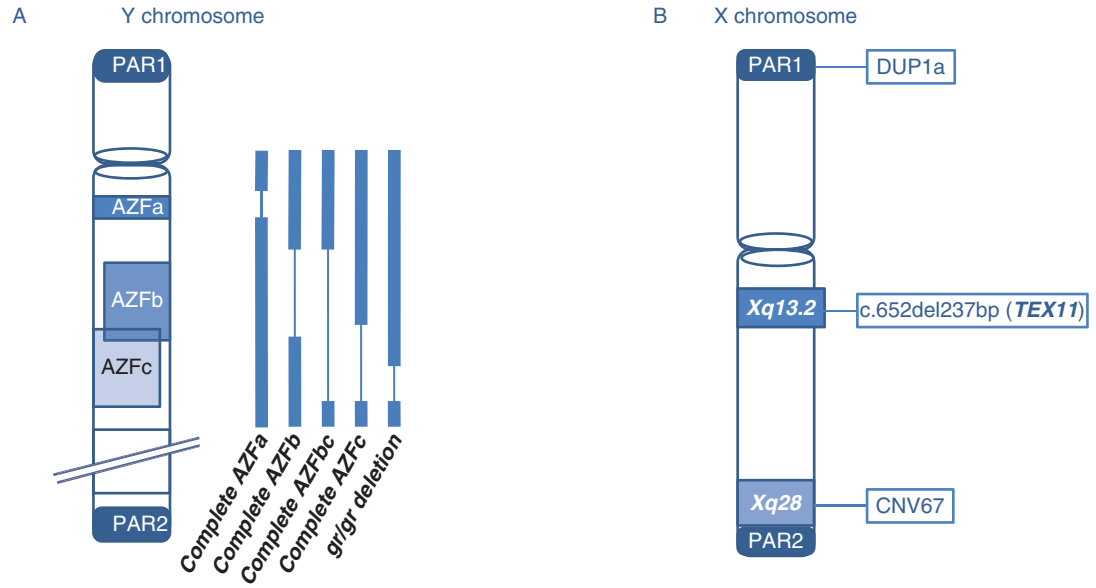
NGS has been recently used with success also for studies of familial cases of azoo/oligozoospermia from Turkey. A novel homozygous mutation in the *NPAS2* gene was reported in three brothers from a consanguineous family, showing variable semen phenotypes ranging from azoospermia to oligozoospermia (Ramasamy *et al*. 2015). Another publication focused on two families: in one case, the most plausible cause for impaired spermatogenesis was a homozygous truncating mutation in *TAF4B;* in the other case, two azoospermic brothers were homozygous for a mutation in the *ZMYND15* gene (Ayhan *et al*. 2014). All these genes are expressed in the testis and are plausible candidates for the observed phenotypes. However, given that the heterozygous carriers of the families are not affected, mutation screening in sporadic NOA patients has limited, if any, diagnostic relevance.

On the contrary, sex chromosomes represent an optimal target in sporadic cases since mutations are in hemizygosis with a potential direct effect on protein function without a compensating effect from a normal allele. Stouffs & Lissens (2012) have reviewed the literature concerning X-linked gene mutations in eight genes. With the exception of the *AR* gene, no other causative mutations/polymorphisms have been described with clinical relevance. Novel data on X chromosome-linked genes derives from recent array-CGH studies (see paragraph below) and the most interesting findings concern genes belonging to the cancer testis antigen (CTA) family (Krausz & Giachini 2012) and to a meiosis genes, TEX11 (Yatsenko *et al*. 2015) (Fig. 1B).

As far as the Y chromosome-linked genes are concerned, studies are limited to deletion analysis rather than intragenic mutation screening, and the only relevant finding concerns the *USP9Y* gene in the AZFa region (Tyler-Smith & Krausz 2009) Deletions affecting this gene have been associated with a variable semen phenotype from azoospermia to normozoospermia, indicating that the gene is more likely a fine tuner than an essential factors for spermatogenesis.

CNVs and male infertility

CNVs are a class of structural variation that may involve complex gains or losses of homologous sequences at multiple sites in the genome. The first genome-wide map of CNVs existing in the human genome showed that these variations cover ~360 Mb, i.e., 12% of the human genome and represent the primary source of inter-individual variability between genomes (Redon *et al*. 2006). Notwithstanding, the gain or loss of DNA sequence can also produce a spectrum of functional effects and human disease phenotypes, by both disrupting gene-coding sequences and affecting region void of genes but involving regulatory elements with an indirect effect on gene transcription. Although the functional consequences of a CNV might be difficult to predict, many CNVs do generate alleles with a clear-cut impact on health and have been associated with a growing number of common complex diseases (Riggs *et al*. 2014). As infertility is indeed a complex disease, it has been hypothesized that certain CNVs may cause defective recombination (especially those mapping to PAR), leading to meiotic failure and the loss of germ cells, or might affect the activity of individual genes important for spermatogenesis. To date, the only CNVs proved to be in a clear-cut cause-effect relationship with spermatogenic impairment are the AZF microdeletions on the Y chromosome (Vogt *et al*. 1996, Krausz *et al*. 2014). Furthermore, the relationship between CNVs and male infertility was also investigated on a larger scale by performing array-CGH on the whole genome (Tüttelmann *et al*. 2011, Stouffs *et al*. 2012, Lopes *et al*. 2013) or at



Deletion	Frequency		
	Patients		Controls
	Azoo	*OAT*	
AZF microdeletions	5–10%	2–5%	0%
gr/gr deletion*	3.5%		0.9%

CNV	CNV type	Frequency	
		Patients (%)	Controls (%)
DUP1a	Duplication	1.4	0
c.652del237bp	Deletion	0.7	0
CNV67	Deletion	1.1	0

Figure 1 Schematic representation of sex chromosome-linked CNVs with clinical relevance. (A) Y chromosome CNVs: the picture illustrates complete AZF microdeletions, a direct cause of impaired spermatogenesis and the gr/gr deletion, an ascertained risk factor for spermatogenic impairment. In the lower table, AZF microdeletions and gr/gr deletion frequencies in patients and controls are reported. Azoo: azoospermic; OAT: oligoasthenoteratozoospermic. * mean frequencies of the gr/gr deletion are relative to the Italian and Spanish populations. (B) X chromosome CNVs: DUP1a (Chianese *et al.* 2014), c.652del237bp in TEX11 (Yatsensko *et al.* 2015) and CNV67 (Lo Giacco *et al.* 2014*a*) are three novel variants with potential clinical implication given their specific association with impaired spermatogenic phenotypes. In the lower table, CNVs type and frequencies in patients and controls are reported. Figure is not in scale.

high resolution on the X chromosome (Krausz *et al.* 2012). The three studies that compared the CNV load between patients and controls all converged on a significantly higher burden of CNVs in men with spermatogenic disturbances (Tüttelmann *et al.* 2011, Krausz *et al.* 2012, Lopes *et al.* 2013). In our study, both the mean number of CNVs/person (mainly dependent on an over-representation of losses) and the mean size/person were significantly increased in the patient group (Krausz *et al.* 2012). In addition, a significantly lower sperm concentration and total sperm count was found in patients with >1 CNV compared to those with ≤ 1 CNV. This excess of X-linked CNVs and DNA loss in patients with reduced sperm count and the significant association between CNV number and sperm count in the infertile group support the existence of a potential link between the observed CNV burden and spermatogenic failure. These conclusions are supported also at the whole genome level, but the CNV burden is especially pronounced on the sex chromosomes (Tüttelmann *et al.* 2011, Lopes *et al.* 2013). More specifically, Tüttelmann *et al.* (2011) reported a significant over-representation of sex-chromosomal CNVs in azoospermic men with Sertoli-cell only (SCO) histology, whereas Lopes *et al.* (2013) in azoo/oligozoospermic men.

Sex chromosomes

Sex chromosomes clearly play an important role in spermatogenesis since they are enriched with genes involved in the development and differentiation of gonads and gametogenesis (Skaletsky *et al.* 2003, Mueller *et al.* 2008, 2013). Given that with the exception of the PAR genes, men are hemizygous for most of the genes located on this chromosome, any *de novo* mutation/CNV might have an immediate impact, since no compensation is provided by another normal allele. Moreover, both chromosomes have accumulated a relevant number of segmental duplications (also called amplicons), which constitute a favorable substrate for CNV formation.

The Y chromosome

The Y chromosome: as already mentioned, Y chromosome microdeletions occurring on the AZF region are the first and thus far the only example of CNVs with clinical significance (Krausz *et al*. 2014). While the complete AZF deletions have been introduced as a routine genetic test for patients with severe OAT and NOA, the role of partial AZFc deletions, i.e., gr/gr deletion, b1/b3, b2/b3 (Repping *et al*. 2003, 2004) has been the object of long-lasting debates (Fig. 1A). Four meta-analyses are available on the gr/gr deletion and all reach significant odds ratios, reporting on average two- to 2.5-fold increased risks of reduced sperm output/infertility (Tüttelmann *et al*. 2007, Visser *et al*. 2009, Navarro-Costa *et al*. 2010, Stouffs *et al*. 2011). In a more recent survey on AZFc deletions in a sample of 20 884 men, Rozen *et al*. (2012) found the gr/gr deletion to be the most common among partial AZFc deletions (2.4% or 1/41 men), as well as that it doubles the risk for impaired spermatogenesis. These data altogether thus confirm the gr/gr deletion as an established significant genetic risk factor for impaired sperm production. The entity of the risk associated with this genetic anomaly varies between populations, reaching the highest OR in Italians, which have a 7.9-fold increased risk for spermatogenic impairment (OR=7.9, 95% CI 1.8–33.8) (Ferlin *et al*. 2005, Giachini *et al*. 2005, 2008). The existence of Y chromosomal haplogroups that constitutively carry the gr/gr deletion, such as the Db2 branch common in Japan and the Q1 haplogroup common in China, indicates that the Y background may modulate the penetrance of this CNV in Asia (Repping *et al*. 2006, Zhang *et al*. 2007). Interestingly, phenotypic variation within European carriers of the Y-chromosomal gr/gr deletion is independent of the Y-chromosomal background (Krausz *et al*. 2009).

Though Y-chromosome microdeletions are directly associated only with spermatogenic failure, concerns have been raised about the potential risk for carriers undergoing assisted reproductive technology to father children affected not only by impaired spermatogenesis but also other conditions such as Turner's syndrome (45,X) and other phenotypic anomalies associated with sex chromosome mosaicism (e.g., ambiguous genitalia) (Patsalis *et al*. 2002, Krausz *et al*. 2014). Furthermore, a recent study (Jorgez *et al*. 2011) reported that 5.4% of men with AZF deletions and a normal karyotype also carried *SHOX* haploinsufficiency. Indeed, this information raised the question about the importance of screening for *SHOX*-linked CNVs in men carrying Y-chromosome microdeletions. Our group performed a large multicenter study in order to evaluate whether such an alarming hypothesis was actually true (Chianese *et al*. 2013). No association was found between Y-chromosome microdeletions and *SHOX* haploinsufficiency, implying that deletion carriers have no augmented risk of *SHOX*-related pathologies (short stature and skeletal anomalies).

The question whether increased gene dosage of the AZFc region may also affect fertility originates from the observation of a limited variation in the copy number of AZFc-linked genes, which strongly indicates a natural selection for the conservation of an 'optimal' copy number by removing exceptionally high or low copy number variants from the population (Repping *et al*. 2006). The *DAZ* gene in the AZFc region is a clear example: about 90% of men carry four *DAZ* copies, which suggests that this is the optimal number required for normal spermatogenesis and that both a reduction and an increase of AZFc gene dosage may have a negative effect. This observation encouraged initially two groups to investigate the clinical consequences of partial AZFc duplications, reaching different conclusions: an association between increased AZFc gene dosage and male infertility was observed in the Han Chinese study (Lin *et al*. 2007), whereas no association could be detected in the Italian study population (Giachini *et al*. 2008). Later on, the effect of AZFc duplications on spermatogenesis was further investigated and again different results were obtained. Ye *et al*. (2013) found a significantly higher frequency of partial duplications in the infertile patients (4.0%) compared to controls (0.7%) in the Chinese-Yi population. Contrastingly, in the analysis by Lo Giacco *et al*. (2014*a*), performed on a study population including prevalently Spanish subjects, AZFc duplications were found at comparable frequencies in patients (4.9%) and controls (3.5%). Seemingly, this discordance reflects mere ethnic differences; therefore, if increased AZFc gene content does play a role in spermatogenic impairment, the effect is probably modulated by population-specific factors.

The X chromosome

The first X chromosome studies were based on the candidate gene approach, and a total of seven X-linked candidate genes have been studied so far (*AR, AKAP, FATE, NXF2, TAF7L, SOX3, USP26*). With the exception of the *AR* gene, no clear-cut causative mutations have been reported and SNPs linked to some of these genes have been the objects of discordant results (Table 1). With the shift of discovery research to high-throughput approaches, researchers were encouraged to apply such technologies to investigate X chromosome-linked CNVs and their role in spermatogenic failure. To date, four groups have employed comparative genomic hybridization (CGH) arrays (Tüttelmann *et al*. 2011, Krausz *et al*. 2012, Stouffs *et al*. 2012, Lopes *et al*. 2013) and three provide information about X-linked CNVs with potential clinical relevance in the etiology of male infertility (Tüttelmann *et al*. 2011, Krausz *et al*. 2012, Lopes *et al*. 2013) (Fig. 1B).

The analysis performed by array-CGH employing a high-resolution (probe distance of 2–4 Kb) X chromosome-specific platform (Krausz *et al*. 2012) allowed the identification of a consistent number of CNVs on the X chromosome, the majority of which (75.3%) were novel. From a clinical standpoint, of particular interest are patient-enriched (significantly more frequent in patients) and patient-specific (not found in controls) CNVs, since genes and regulatory elements within or nearby these regions presumably have a higher probability of being implicated in spermatogenic failure. Although there are some partially overlapping findings regarding the X chromosome-linked CNVs between the three studies (Tüttelmann *et al*. 2011, Krausz *et al*. 2012, Lopes *et al*. 2013), differences in the resolution of the arrays may explain the lack of complete overlaps. By performing a comparison between the raw data of the three studies we observed a few interesting overlapping CNVs. Three patient-specific CNVs – DUP1a, DUP55 and DUP60 – detected in the study by Krausz *et al*. (2012) were also found by Tüttelmann *et al*. (2011) in men affected by SCOS. The comparison with data by Lopes *et al*. (2013) also shows an overlap of a recurrent deletion detected in their study at a significantly higher frequency in patients compared to controls and two patient-specific CNVs, CNV30 (gain) and CNV31 (loss), identified in the Krausz' study. When comparing patient-specific CNVs detected in the study by Tüttelmann *et al*. (2011), the loss nssv1496532 overlaps with CNV69, which was found significantly more frequent in patients than controls in the Krausz' study. One gain on Xq22.2 (Lopes *et al*. 2013) overlapped with the private duplication nssv1499049 found in an oligozoospermic man in Tüttelmann's study. It is worth noting that this duplication intersects a number of genes with specific or exclusive expression in the testis (*H2BFWT, H2BFXP* and *H2BFM*). No CNVs were found to be common to all three studies. In the light of these comparisons, DUP1a, CNV69 and the nssv1499049 are promising variants, since their potential involvement in spermatogenic impairment was reported by more than one study.

In fact, the two variants DUP1a and CNV69 were objects of large follow-up studies, together with other recurrent deletions, CNV67 and CNV64 (Chianese *et al*. 2014, Lo Giacco *et al*. 2014*b*). The first study analyzed three recurrent deletions (frequency >1%) in a large case–control setting ($n=1255$) for their exclusive (CNV67) and prevalent (CNV64 and CNV69) presence in patients. For instance, deletion carriers displayed a higher probability of having impaired spermatogenesis (OR=1.9 and 2.2 for CNV64 and CNV69 respectively) as well as sperm concentration and total motile sperm number was lower in carriers compared to non-carriers The most interesting deletion was CNV67 because it was exclusively found in patients with a frequency of 1.1% ($P<0.01$) and is likely to involve the *MAGE9A* gene – a CTA family member – and/or its regulatory elements (Lo Giacco *et al*. 2014*b*). Similarly, a follow-up study was performed on five selected gains (DUP1A, DUP5, DUP20, DUP26 and DUP40), which include, or are in close proximity to, genes with testis-specific expression and potential implication in spermatogenesis (Chianese *et al*. 2014). While four of the five CNVs (DUP5, DUP20, DUP26 and DUP40) did not individually reach statistical significance, they remained patient-specific. DUP1A, instead, was found exclusively and at a significantly higher frequency in patients. This gain fully duplicates a long non-coding RNA (*LINC00685*) that potentially acts as a negative regulator of a gene with potential role in spermatogenesis, *PPP2R3B*; according to our hypothesis, the mechanism by which DUP1A could lead to spermatogenic failure is a misbalanced ratio of the *PPP2R3B* and its antisense, causing a decrease in *PPP2R3B* transcription in the developing germ cells (Chianese *et al*. 2014). Our data together with the identification of two SCOS patients with a duplication disrupting the *PPP2R3B* gene (Tüttelmann *et al*. 2011) indicate that CNVs mapping into this region and affecting either *PPP2R3B* or the long non-coding RNA (*LINC00685*) are good mutational targets for future case–control studies.

Lastly, a recent study proved the implication of the *TEX11* gene in meiotic arrest and azoospermia (Yatsenko *et al*. 2015). The study population included a total of 289 patients with different testis histology (63 with SCOS, 33 with meiotic arrest and 193 with mixed testicular atrophy) and 384 normozoospermic controls. With the use of an X-chromosome high-resolution GCH microarray, they firstly analyzed 15 azoospermic men and found that a patient with mixed atrophy carried a 91-KB deletion (c.652del237bp) encompassing exons 10, 11 and 12 of *TEX11*. Further Sanger sequencing in the rest of the patients allowed detecting that another man with meiotic arrest carried the same deletion c.652del237bp, which was confirmed by array-CGH validation; moreover, they found five patients with either meiotic arrest or mixed testicular atrophy carrying missense mutations in *TEX11*. None of the controls carried any of these variants. Finally, the finding of *TEX11* mutations in 2.4% ($n=7/289$) of patients, of which 15% ($n=5/33$) suffered from meiotic arrest and 1% ($n=2/193$) had a mixed testicular atrophy, supports the importance of this gene for normal spermatogenesis.

Autosomes

Whole-genome approaches allowed providing data also on the potential role of autosome-linked CNVs in relation to different semen phenotypes (Tüttelmann *et al*. 2011, Stouffs *et al*. 2012, Lopes *et al*. 2013). The first study reported eight autosomal rearrangements (involving chromosomes 1, 2, 3, 5, 12, 15, 16, 17) potentially linked to fertility problems, as they were not detected in normozoospermic controls (Stouffs *et al*. 2012).

The second study reported recurrent and patient-specific autosomal CNVs potentially associated with oligozoospermia ($n=11$) and with SCOS ($n=4$), also reporting a list of genes intersecting the CNVs and with potential involvement in the spermatogenic phenotype. Finally, after assaying genome-wide SNPs and CNVs, the third study estimated that rare autosomal deletions multiplicatively change a man's risk of disease by 10% (OR 1.10 (1.04–1.16), $P<2\times10^{-3}$). The same authors observed five deletions (ranging in size from 54 kb to over 2 Mb) of the autosomal *DMRT1* gene in four cases of azoospermia and one in normozoospermia. Despite the normozoospermic deletion carrier, statistical analysis based on the comparison of all patients versus 7000 controls lead to a significant association with impaired sperm production. Given the low frequency of this mutation and the wide range of associated phenotype, it remains difficult to include the testing for DMRT1-linked CNVs in the routine diagnostic workup.

The comparison between the three studies shows some overlapping findings. When comparing the CNVs detected by Stouffs *et al.* (2012) with the raw data deposited in dbVar by Tüttelmann *et al.* (2011), five overlapping loci can be observed on chromosomes 1, 5, 15, 16 and 17, but only those related to chromosome 1 and 16 results are patient-specific in both studies. The first locus on chromosome 1 shares a 46 kb-span overlap with the gain nssv1495850 reported in an oligozoospermic man in Tüttelmann's study. The other locus on chromosome 16 overlaps with both gains and losses from Tüttelmann's study; interestingly, gains are found in both patients and controls, whereas the reciprocal losses were exclusively detected in OAT patients. When comparing the Lopes' and the Tüttelmann's study, one overlap is reported on chromosome 8: at this locus, Tuttelmann *et al.* identified a deletion in an azoospermic man and another with a duplication, intersecting the *PLEC1* and *MIR661* genes, whereas Lopes *et al.* identified a duplication in an oligozoospermic man affecting the same genes. No CNVs were observed to be common to all three studies.

Summary and future directions

Male infertility is a multifactorial complex disease with highly heterogeneous phenotypic representation. The wide range of quantitative and qualitative impairments can be caused by several acquired and congenital factors, including genetic/epigenetic anomalies. Despite a 10-year effort, research was largely unsuccessful in identifying recurrent genetic factors with potential clinical application. The armamentarium of diagnostic tests has been implemented only by the screening for Y chromosome-linked gr/gr deletion in those populations for which robust and consistent data with risk estimate are available. Much expectation was given to genome-wide SNP arrays, based on the analysis of common variants, but no overlapping SNPs have been identified between different studies. Meta-analyses have been able to demonstrate significant association only for a few SNPs, conferring generally weak predisposition to infertility. According to a few observations, common SNPs with significant but low effect size may eventually lead to impaired spermatogenic efficiency if they are present contemporarily in the same individual (Aston *et al.* 2010, Kosova *et al.* 2012). On the other hand, it is clearly demonstrated by both SNP and array-CGH, that there is a rare variant burden in men with impaired spermatogenesis, which is especially relevant concerning CNVs. Whether this phenomenon is an expression of a more generalized genomic instability is still an open question. Epidemiological observations indicating lower life expectancy and higher morbidity in infertile men (Jensen *et al.* 2009, Salonia *et al.* 2009, Eisenberg *et al.* 2014) are suggestive for such a potential relationship.

It has been predicted that more than 2000 genes (housekeeping and specific germ cell genes) are involved in spermatogenesis (Hochstenbach & Hackstein, 2000) and mutation in these genes may act directly or through gene-environmental interaction. In the era of NGS we expect to expand our diagnostic skills, since mutations in several hundred of genes can potentially lead to infertility and each of them is likely responsible for only a small fraction of cases. Exome analysis is predicted to be successful especially for descendants of consanguineous families and familial cases of infertility. Concerning sporadic oligo/azoospermia, the situation is more complex and, since the infertile trait undergoes negative selection, at least two scenarios can be predicted. On one hand, there is a possibility that rare or *de novo* large-effect mutations are involved in these pathological conditions; in this regard, the X chromosome represents one of the most exciting future targets for both its enrichment in genes involved in spermatogenesis and its hemizygous state in males, which implies a direct effect of a damaging mutation. On the other hand, an alternative pathogenic mechanism can be related to a synergistic effect of multiple heterozygous mutations in genes involved in the same biological pathway. In this regard, system biology, which allows unrevealing possible gene interactions and common biological pathways, will provide an informative tool for NGS data interpretation. Although these novel approaches will certainly help discover 'hidden' genetic factors, a more comprehensive picture of the etiopathogenesis of idiopathic male infertility will only be achieved by a parallel investigation of the complex world of gene environmental interaction and epigenetics.

Supplementary data

This is linked to the online version of the paper at http://dx.doi.org/10.1530/REP-15-0261.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of this review.

Funding

This work was supported by the following funding agencies: European Commission, Reproductive Biology Early Research Training (REPROTRAIN, Project Number: 289880) and the Spanish Ministry of Health (FIS grant PI14/01250).

Acknowledgements

The authors thank Gianni Forti (University of Florence) and Esperança Marti Salas (Fundació Puigvert) for their continuous support to research in genetics of male infertility. We also thank the following funding agencies: European Commission, Reproductive Biology Early Research Training (REPROTRAIN, Project Number: 289880), IBSA Foundation for Scientific Research to CC and the Spanish Ministry of Health (FIS grant PI14/01250).

References

Abbas M, Srivastava K, Imran M & Banerjee M 2014 Association of CYP1A1 gene variants rs4646903 (T>C) and rs1048943 (A>G) with cervical cancer in a North Indian population. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* **176** 68–74. (doi:10.1016/j.ejogrb.2014.02.036)

Aston KI & Carrell DT 2009 Genome-wide study of single-nucleotide polymorphisms associated with azoospermia and severe oligozoospermia. *Journal of Andrology* **30** 711–725. (doi:10.2164/jandrol.109.007971)

Aston KI, Krausz C, Laface I, Ruiz-Castané E & Carrell DT 2010 Evaluation of 172 candidate polymorphisms for association with oligozoospermia or azoospermia in a large cohort of men of European descent. *Human Reproduction* **25** 1383–1397. (doi:10.1093/humrep/deq081)

Atanassova N, McKinnell C, Walker M, Turner KJ, Fisher JS, Morley M, Millar MR, Groome NP & Sharpe RM 1999 Permanent effects of neonatal estrogen exposure in rats on reproductive hormone levels, sertoli cell number, and the efficiency of spermatogenesis in adulthood. *Endocrinology* **140** 5364–5373.

Ayhan Ö, Balkan M, Guven A, Hazan R, Atar M, Tok A & Tolun A 2014 Truncating mutations in TAF4B and ZMYND15 causing recessive azoospermia. *Journal of Medical Genetics* **51** 239–244. (doi:10.1136/jmedgenet-2013-102102)

Bashamboo A, Ferraz-de-Souza B, Lourenço D, Lin L, Sebire NJ, Montjean D, Bignon-Topalovic J, Mandelbaum J, Siffroi JP, Christin-Maitre S *et al.* 2010 Human male infertility associated with mutations in NR5A1 encoding steroidogenic factor 1. *American Journal of Human Genetics* **8** 505–512. (doi:10.1016/j.ajhg.2010.09.009)

Becherini L, Guarducci E, Degl'Innocenti S, Rotondi M, Forti G & Krausz C 2004 DAZL polymorphisms and susceptibility to spermatogenic failure: an example of remarkable ethnic differences. *International Journal of Andrology* **27** 375–381. (doi:10.1111/j.1365-2605.2004.00520.x)

Bianco B, Ghirelli-Filho M, Cavalheiro CM, Cavalcanti V, Peluso C, Gava MM, Glina S, Christofolini DM & Barbosa CP 2013 Variants in endothelial nitric oxide synthase (eNOS) gene in idiopathic infertile Brazilian men. *Gene* **519** 13–17. (doi:10.1016/j.gene.2013.02.001)

Brooke G, Holbrook JD, Brown MH & Barclay AN 2004 Human lymphocytes interact directly with CD47 through a novel member of the signal regulatory protein (SIRP) family. *Journal of Immunology* **173** 2562–2570. (doi:10.4049/jimmunol.173.4.2562)

Buldreghini E, Mahfouz RZ, Vignini A, Mazzanti L, Ricciardo-Lamonica G, Lenzi A, Agarwal A & Balercia G 2010 Single nucleotide polymorphism (SNP) of the endothelial nitric oxide synthase (eNOS) gene (Glu298Asp variant) in infertile men with asthenozoospermia. *Journal of Andrology* **31** 482–488. (doi:10.2164/jandrol.109.008979)

Chen P, Wang X, Xu C, Xiao H, Zhang WH, Wang XH & Zhang XH 2015 Association of polymorphisms of A260G and A386G in DAZL gene with male infertility: a meta-analysis and systemic review. *Asian Journal of Andrology* [in press]. (doi:10.4103/1008-682X.153542)

Chianese C, oGiacco D, Tüttelmann F, Ferlin A, Ntostis P, Vinci S, Balercia G, Ars E, Ruiz-Castañé E, Giglio S, Ruiz-Castañé E, Giglio S *et al.* 2013 Y-chromosome microdeletions are not associated with SHOX haploinsufficiency. *Human Reproduction* **28** 3155–3160. (doi:10.1093/humrep/det322)

Chianese C, Gunning AC, Giachini C, Daguin F, Balercia G, Ars E, Lo Giacco D, Ruiz-Castañé E, Forti G, Krausz C, Ruiz-Castañé E, Forti G & Krausz C 2014 X chromosome-linked CNVs in male infertility: discovery of overall duplication load and recurrent, patient-specific gains with potential clinical relevance. *PLoS ONE* **9** e97746. (doi:10.1371/journal.pone.0097746)

Chihara M, Yoshihara K, Ishiguro T, Yokota Y, Adachi S, Okada H, Kashima K, Sato T, Tanaka A, Tanaka K *et al.* 2015 Susceptibility to male infertility: replication study in Japanese men looking for an association with four GWAS-derived loci identified in European men. *Journal of Assisted Reproduction and Genetics* **32** 903–908. (doi:10.1007/s10815-015-0468-4)

Choi Y, Jeon S, Choi M, Lee M, Park M, Lee DR, Jun K-Y, Kwon Y, Lee O-H, Song S-H *et al.* 2010 Mutations in SOHLH1 gene associate with nonobstructive azoospermia. *Human Mutation* **31** 788–793. (doi:10.1002/humu.21264)

Davis-Dao CA, Tuazon ED, Sokol RZ & Cortessis VK 2007 Male infertility and variation in CAG repeat length in the androgen receptor gene: a meta-analysis. *Journal of Clinical Endocrinology and Metabolism* **92** 4319–4326. (doi:10.1210/jc.2007-1110)

Dosenko VE, Zagoriy VY, Haytovich NV, Gordok OA & Moibenko AA 2006 Allelic polymorphism of endothelial NO-synthase gene and its functional manifestations. *Acta Biochimica Polonica* **53** 299–302.

Edelmann W, Cohen PE, Kneitz B, Winand N, Lia M, Heyer J, Kolodner R, Pollard JW & Kucherlapati R 1999 Mammalian MutS homologue 5 is required for chromosome pairing in meiosis. *Nature Genetics* **21** 123–127. (doi:10.1038/5075)

Eisenberg ML, Li S, Behr B, Cullen MR, Galusha D, Lamb DJ & Lipshultz LI 2014 Semen quality, infertility and mortality in the USA. *Human Reproduction* **29** 1567–1574. (doi:10.1093/humrep/deu106)

Ferlin A, Tessari A, Ganz F, Marchina E, Barlati S, Garolla A, Engl B & Foresta C 2005 Association of partial AZFc region deletions with spermatogenic impairment and male infertility. *Journal of Medical Genetics* **42** 209–213. (doi:10.1136/jmg.2004.025833)

Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, Boers GJ, den Heijer M, Kluijtmans LA, van den Heuvel LP *et al.* 1995 A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nature Genetics* **10** 111–113. (doi:10.1038/ng0595-111)

Ge YZ, Xu LW, Jia RP, Xu Z, Li WC, Wu R, Liao S, Gao F, Tan SJ, Song Q *et al.* 2014 Association of polymorphisms in estrogen receptors (ESR1 and ESR2) with male infertility: a meta-analysis and systematic review. *Journal of Assisted Reproduction and Genetics* **31** 601–611. (doi:10.1007/s10815-014-0212-5)

Giachini C, Guarducci E, Longepied G, Degl'Innocenti S, Becherini L, Forti G, Mitchell MJ & Krausz C 2005 The gr/gr deletion(s): a new genetic test in male infertility? *Journal of Medical Genetics* **42** 497–502. (doi:10.1136/jmg.2004.028191)

Giachini C, Laface I, Guarducci E, Balercia G, Forti G & Krausz C 2008 Partial AZFc deletions and duplications: clinical correlates in the Italian population. *Human Genetics* **124** 399–410. (doi:10.1007/s00439-008-0561-1)

Gineitis AA, Zalenskaya IA, Yau PM, Bradbury EM & Zalensky AO 2001 Human sperm telomere – binding complex involves histone H2B and secures telomere membrane. *Journal of Cell Biology* **151** 1591–1598. (doi:10.1083/jcb.151.7.1591)

Gong M, Dong W, He T, Shi Z, Huang G, Ren R, Huang S, Qiu S & Yuan R 2015 MTHFR 677C>T polymorphism Increases the male infertility risk: a meta-analysis involving 26 studies. *PLoS ONE* **10** e0121147. (doi:10.1371/journal.pone.0121147)

Guéant JL, Romano A, Cornejo-Garcia JA, Oussalah A, Chery C, BlancaLópez N, Guéant-Rodriguez RM, Gaeta F, Rouyer P, Josse T *et al.* 2015 HLA-DRA variants predict penicillin allergy in genome-wide fine-mapping genotyping. *Journal of Allergy and Clinical Immunology* **135** 253–259. (doi:10.1016/j.jaci.2014.07.047)

Gupta N, Gupta S, Dama M, David A, Khanna G, Khanna A & Rajender S 2011 Strong association of 677 C>T substitution in the MTHFR gene with male infertility – a study on an indian population and a meta-analysis. *PLoS ONE* **6** e22277. (doi:10.1371/journal.pone.0022277)

Hess RA 2003 Estrogen in the adult male reproductive tract: a review. *Reproductive Biology and Endocrinology* **1** 52.

Hochstenbach R & Hackstein JH 2000 The comparative genetics of human spermatogenesis: clues from flies and other model organisms. *Results and Problems in Cell Differentiation* **28** 271–298.

Hu Z, Xia Y, Guo X, Dai J, Li H, Hu H, Jiang Y, Lu F, Wu Y, Yang X *et al.* 2012 A genome-wide association study in Chinese men identifies three risk loci for non-obstructive azoospermia. *Nature Genetics* **44** 183–186. (doi:10.1038/ng.1040)

Hu Z, Li Z, Yu J, Tong C, Lin Y, Guo X, Lu F, Dong J, Xia Y, Wen Y *et al.* 2014 Association analysis identifies new risk loci for non-obstructive azoospermia in Chinese men. *Nature Communications* **5** 3857.

Jensen TK, Jacobsen R, Christensen K, Nielsen NC & Bostofte E 2009 Good semen quality and life expectancy: a cohort study of 43,277 men. *American Journal of Epidemiology* **170** 559–565. (doi:10.1093/aje/kwp168)

Ji G, Long Y, Zhou Y, Huang C, Gu A & Wang X 2012 Common variants in mismatch repair genes associated with increased risk of sperm DNA damage and male infertility. *BMC Medicine* **10** 49. (doi:10.1186/1741-7015-10-49)

Jinam TA, Nakaoka H, Hosomichi K, Mitsunaga S, Okada H, Tanaka A, Tanaka K & Inoue I 2013 HLA-DPB1*04:01 allele is associated with non-obstructive azoospermia in Japanese patients. *Human Genetics* **132** 1405–1411. (doi:10.1007/s00439-013-1347-7)

Jorgez CJ, Weedin JW, Sahin A, Tannour-Louet M, Han S, Bournat JC, Mielnik A, Cheung SW, Nangia AK, Schlegel PN *et al.* 2011 Aberrations in pseudoautosomal regions (PARs) found in infertile men with Y-chromosome microdeletions. *Journal of Clinical Endocrinology and Metabolism* **96** E674–E679. (doi:10.1210/jc.2010-2018)

Kosova G, Scott NM, Niederberger C, Prins GS & Ober C 2012 Genome-wide association study identifies candidate genes for male fertility traits in humans. *American Journal of Human Genetics* **90** 950–961. (doi:10.1016/j.ajhg.2012.04.016)

Krausz C 2011 Male infertility: pathogenesis and clinical diagnosis. *Best Practice & Research. Clinical Endocrinology & Metabolism* **25** 271–285. (doi:10.1016/j.beem.2010.08.006)

Krausz C, Giachini C, Xue Y, O'Bryan MK, Gromoll J, Rajpert-de Meyts E, Oliva R, Aknin-Seifer I, Erdei E, Jorgensen N *et al.* 2009 Phenotypic variation within European carriers of the Y-chromosomal gr/gr deletion is independent of Y chromosomal background. *Journal of Medical Genetics* **46** 21–31. (doi:10.1136/jmg.2008.059915)

Krausz C, Giachini C, Lo Giacco D, Daguin F, Chianese C, Ars E, Ruiz-Castane E, Forti G & Rossi E 2012 High resolution X chromosome-specific array-CGH detects new CNVs in infertile males. *PLoS ONE* **7** e44887. (doi:10.1371/journal.pone.0044887)

Krausz C, Hoefsloot L, Simoni M & Tüttelmann F 2014 EAA/EMQN best practice guidelines for molecular diagnosis of Y-chromosomal microdeletions: state-of-the-art 2013. *Andrology* **2** 5–19. (doi:10.1111/j.2047-2927.2013.00173.x)

Kuo YC, Lin YH, Chen HI, Wang YY, Chiou YW, Lin HH, Pan HA, Wu CM, Su SM, Hsu CC *et al.* 2012 SEPT12 mutations cause male infertility with defective sperm annulus. *Human Mutation* **33** 710–719. (doi:10.1002/humu.22028)

Kusz-Zamelczyk K, Sajek M, Spik A, Glazar R, Jędrzejczak P, Latos-Bieleńska A, Kotecki M, Pawelczyk L & Jaruzelska J 2013 Mutations of NANOS1, a human homologue of the Drosophila morphogen, are associated with a lack of germ cells in testes or severe oligo-astheno-teratozoospermia. *Journal of Medical Genetics* **50** 187–193. (doi:10.1136/jmedgenet-2012-101230)

Lee NPY & Cheng CY 2008 Nitric oxide and cyclic nucleotides their roles in junction dynamics and spermatogenesis. *Oxidative Medicine and Cellular Longevity* **1** 25–32. (doi:10.4161/oxim.1.1.6856)

Lee J, Park HS, Kim HH, Yun YJ, Lee DR & Lee S 2009 Functional polymorphism in H2BFWT-5′UTR is associated with susceptibility to male infertility. *Journal of Cellular and Molecular Medicine* **13** 1942–1951. (doi:10.1111/j.1582-4934.2009.00830.x)

Li Z, Huang Y, Li H, Hu J, Liu X, Jiang T, Sun G, Tang A, Sun X, Qian W *et al.* 2015 Excess of rare variants in genes that are key epigenetic regulators of spermatogenesis in the patients with non-obstructive azoospermia. *Scientific Reports* **5** article 8785. (doi:10.1038/srep08785)

Lin YW, Hsu LCL, Kuo PL, Huang WJ, Chiang HS, Yeh SD, Hsu TY, Yu YH, Hsiao KN, Cantor RM *et al.* 2007 Partial duplication at AZFc on the Y chromosome is a risk factor for impaired spermatogenesis in Han Chinese in Taiwan. *Human Mutation* **28** 486–494. (doi:10.1002/humu.20473)

Lin X, Deng FY, Mo XB, Wu LF & Lei S-F 2015 Functional relevance for multiple sclerosis-associated genetic variants. *Immunogenetics* **67** 7–14. (doi:10.1007/s00251-014-0803-4)

Lo Giacco D, Chianese C, Sánchez-Curbelo J, Bassas L, Ruiz P, Rajmil O, Sarquella J, Vives A, Ruiz-Castañé E & Oliva R 2014*a* Clinical relevance of Y-linked CNV screening in male infertility: new insights based on the 8-year experience of a diagnostic genetic laboratory. *European Journal of Human Genetics* **22** 754–761. (doi:10.1038/ejhg.2013.253)

Lo Giacco D, Chianese C, Ars E, Ruiz-Castane E, Forti G & Krausz C 2014*b* Recurrent X chromosome-linked deletions: discovery of new genetic factors in male infertility. *Journal of Medical Genetics* **51** 340–344. (doi:10.1136/jmedgenet-2013-101988)

Lopes AM, Aston KI, Thompson E, Carvalho F, Gonçalves J, Huang N, Matthiesen R, Noordam MJ, Quintela I, Ramu A *et al.* 2013 Human spermatogenic failure purges deleterious mutation load from the autosomes and both sex chromosomes, including the gene DMRT1. *PLoS Genetics* **9** e1003349. (doi:10.1371/journal.pgen.1003349)

Lu C, Xu M, Wang R, Qin Y, Wang Y, Wu W, Song L, Wang S, Shen H, Sha J *et al.* 2014 Pathogenic variants screening in five non-obstructive azoospermia-associated genes. *Molecular Human Reproduction* **20** 178–183. (doi:10.1093/molehr/gat071)

Luo H, Li H, Yao N, Hu L & He T 2014 Association between 3801T>C polymorphism of CYP1A1 and idiopathic male infertility risk: a systematic review and meta-analysis. *PLoS ONE* **9** e86649. (doi:10.1371/journal.pone.0086649)

Manfredi S, Federici C, Picano E, Botto N, Rizza A & Andreassi MG 2007 GSTM1, GSTT1 and CYP1A1 detoxification gene polymorphisms and susceptibility to smoking-related coronary artery disease: a case-only study. *Mutation Research* **621** 106–112. (doi:10.1016/j.mrfmmm.2007.02.014)

Mcmanus ME, Burgess WM, Veronese ME, Huggett A, Quattrochi LC & Tukey RH 1990 Metabolism of 2-acetylaminofluorene and benzo (a) pyrene and activation of food- derived heterocyclic amine mutagens by human cytochromes P-450. *Cancer Research* **50** 3367–3376.

Mou L, Wang Y, Li H, Huang Y, Jiang T, Huang W, Li Z, Chen J, Xie J, Liu Y *et al.* 2013 A dominant-negative mutation of HSF2 associated with idiopathic azoospermia. *Human Genetics* **132** 159–165. (doi:10.1007/s00439-012-1234-7)

Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC & Turner JM 2008 The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nature Genetics* **40** 794–799. (doi:10.1038/ng.126)

Mueller JL, Skaletsky H, Brown LG, Zaghlul S, Rock S, Graves T, Auger K, Warren WC, Wilson RK & Page DC 2013 Independent specialization of the human and mouse X chromosomes for the male germ line. *Nature Genetics* **45** 1083–1087. (doi:10.1038/ng.2705)

Navarro-Costa P, Gonçalves J & Plancha CE 2010 The AZFc region of the Y chromosome: at the crossroads between genetic diversity and male infertility. *Human Reproduction Update* **16** 525–542. (doi:10.1093/humupd/dmq005)

Neefjes J, Jongsma MLM, Paul P & Bakke O 2011 Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews. Immunology* **11** 823–836. (doi:10.1038/nri3084)

Nenonen HA, Giwercman A, Hallengren E & Giwercman YL 2011 Non-linear association between androgen receptor CAG repeat length and risk of male subfertility – a meta-analysis. *International Journal of Andrology* **34** 327–332. (doi:10.1111/j.1365-2605.2010.01084.x)

Nuti F & Krausz C 2008 Gene polymorphisms/mutations relevant to abnormal spermatogenesis. *Reproductive Biomedicine Online* **16** 504–513. (doi:10.1016/S1472-6483(10)60457-9)

Patsalis PC, Sismani C, Quintana-Murci L, Taleb-Bekkouche F, Krausz C & McElreavey K 2002 Effects of transmission of Y chromosome AZFc deletions. *Lancet* **19** 1222–1224. (doi:10.1016/S0140-6736(02)11248-7)

Plaseski T, Noveski P, Popeska Z, Efremov GD & Plaseska-Karanfilska D 2012 Association study of single-nucleotide polymorphisms in FAS-LG,JMJDIA, LOC203413, TEX15, BRDT, OR2W3, INSR, and TAS2R38 genes with male infertility. *Journal of Andrology* **33** 675–683. (doi:10.2164/jandrol.111.013995)

van der Put NM, Gabreëls F, Stevens EM, Smeitink JA, Trijbels FJ, Eskes TK, van der Heuvel LP & Blom HJ 1998 A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects? *American Journal of Human Genetics* **62** 1044–1051. (doi:10.1086/301825)

Ramasamy R, Baırcıoğklu ME, Cengiz C, Karaca E, Scovell J, Jhangiani SN, Akdemir ZC, Bainbridge M, Yu Y, Huff C *et al.* 2015 Whole-exome sequencing identifies novel homozygous mutation in NPAS2 in family with nonobstructive azoospermia. *Fertility and Sterility* **104** 286–291. (doi:10.1016/j.fertnstert.2015.04.001)

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W *et al.* 2006 Global variation in copy number in the human genome. *Nature* **444** 444–454. (doi:10.1038/nature05329)

Repping S, Skaletsky H, Brown L, van Daalen SKM, Korver CM, Pyntikova T, Kuroda-Kawaguchi T, de Vries JWA, Oates RD, Silber S *et al.* 2003 Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection. *Nature Genetics* **35** 247–251. (doi:10.1038/ng1250)

Repping S, van Daalen SKM, Korver CM, Brown LG, Marszalek JD, Gianotten J, Oates RD, Silber S, van der Veen F, Page DC *et al.* 2004 A family of human Y chromosomes has dispersed throughout northern Eurasia despite a 1.8-Mb deletion in the azoospermia factor c region. *Genomics* **83** 1046–1052. (doi:10.1016/j.ygeno.2003.12.018)

Repping S, van Daalen SKM, Brown LG, Korver CM, Lange J, Marszalek JD, Pyntikova T, van der Veen F, Skaletsky H, Page DC *et al.* 2006 High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nature Genetics* **38** 463–467. (doi:10.1038/ng1754)

Riggs ER, Ledbetter DH & Martin CL 2014 Genomic variation: lessons learned from whole-genome CNV analysis. *Current Genetic Medicine Reports* **18** 146–150. (doi:10.1007/s40142-014-0048-4)

Rozen SG, Marszalek JD, Irenze K, Skaletsky H, Brown LG, Oates RD, Silber SJ, Ardlie K & Page DC 2012 AZFc deletions and spermatogenic failure: a population-based survey of 20,000 Y chromosomes. *American Journal of Human Genetics* **91** 890–896. (doi:10.1016/j.ajhg.2012.09.003)

Safarinejad MR, Shafiei N & Safarinejad S 2010 The role of endothelial nitric oxide synthase (eNOS) T-786C, G894T, and 4a/b gene polymorphisms in the risk of idiopathic male infertility. *Molecular Reproduction and Development* **77** 720–727. (doi:10.1002/mrd.21210)

Salnikova LE, Belopolskaya OB, Zelinskaya NI & Rubanovich AV 2013 The potential effect of gender in CYP1A1 and GSTM1 genotype-specific associations with pediatric brain tumor. *Tumour Biology* **34** 2709–2719. (doi:10.1007/s13277-013-0823-y)

Salonia A, Matloob R, Gallina A, Abdollah F, Saccà A, Briganti A, Suardi N, Colombo R, Rocchini L, Guazzoni G *et al.* 2009 Are infertile men less healthy than fertile men? Results of a prospective case–control survey *European Urology* **56** 1025–1031. (doi:10.1016/j.eururo.2009.03.001)

Sato Y, Jinam T, Iwamoto T, Yamauchi A, Imoto I, Inoue I & Tajima A 2013 Replication study and meta-analysis of human nonobstructive azoospermia in Japanese populations 1. *Biology of Reproduction* **88** 1–4. (doi:10.1095/biolreprod.112.105957)

Sato Y, Tajima A, Tsunematsu K, Nozawa S, Yoshiike M, Koh E, Kanaya J, Namiki M, Matsumiya K, Tsujimura A *et al.* 2015 An association study of four candidate loci for human male fertility traits with male infertility. *Human Reproduction* **30** 1510–1514. (doi:10.1093/humrep/dev088)

Shen O, Liu R, Wu W, Yu L & Wang X 2012 Association of the methylenetetrahydrofolate reductase gene A1298C polymorphism with male infertility: a meta-analysis. *Annals of Human Genetics* **76** 25–32. (doi:10.1111/j.1469-1809.2011.00691.x)

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T *et al.* 2003 The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **19** 825–837. (doi:10.1038/nature01722)

Smith JG & Newton-Cheh C 2015 Genome-wide association studies of late-onset cardiovascular disease. *Journal of Molecular and Cellular Cardiology* **83** 131–141. (doi:10.1016/j.yjmcc.2015.04.004)

Song P, Zou S, Chen T, Chen J, Wang Y, Yang J, Song Z, iang H, Shi H, Huang Y *et al.* 2015 Endothelial nitric oxide synthase (eNOS) T-786C, 4a4b, and G894T polymorphisms and male infertility: study for idiopathic asthenozoospermia and meta-analysis. *Biology of Reproduction* **92** 38. (doi:10.1095/biolreprod.114.123240)

Stouffs K & Lissens W 2012 X chromosomal mutations and spermatogenic failure. *Biochimica et Biophysica Acta* **1822** 1864–1872. (doi:10.1016/j.bbadis.2012.05.012)

Stouffs K, Lissens W, Tournaye H & Haentjens P 2011 What about gr/gr deletions and male infertility? Systematic review and meta-analysis *Human Reproduction Update* **17** 197–209. (doi:10.1093/humupd/dmq046)

Stouffs K, Vandermaelen D, Massart A, Menten B, Vergult S, Tournaye H & Lissens W 2012 Array comparative genomic hybridization in male infertility. *Human Reproduction* **27** 921–929. (doi:10.1093/humrep/der440)

Takasaki N, Tachibana K, Ogasawara S, Matsuzaki H, Hagiuda J, Ishikawa H, Mochida K, Inoue K, Ogonuki N, Ogura A *et al.* 2014 A heterozygous mutation of GALNTL5 affects male infertility with impairment of sperm motility. *PNAS* **111** 1120–1125. (doi:10.1073/pnas.1310777111)

Tsujimura A, Ota M, Katsuyama Y, Sada M, Miura H, Matsumiya K, Gotoh R, Nakatani T, Okuyama A & Takahara S 2002 Susceptibility gene for non-obstructive azoospermia located near HLA-DR and -DQ loci in the HLA class II region. *Human Genetics* **110** 192–197. (doi:10.1007/s00439-001-0657-3)

Tu W, Liu Y, Shen Y, an Y, Wang X, Yang D, Li L, Ma Y, Tao D, Zhang S *et al.* 2014 Genome-wide loci linked to non-obstructive azoospermia susceptibility may be independent of reduced sperm production in males with normozoospermia. *Biology of Reproduction* **92** 41. (doi:10.1095/biolreprod.114.125237)

Tut TG, Ghadessy FJ, Trifiro MA, Pinsky L & Yong EL 1997 Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility. *Journal of Clinical Endocrinology and Metabolism* **82** 3777–3782.

Tüttelmann F, Rajpert-DeMeyts E, Nieschlag E & Simoni M 2007 Gene polymorphisms and male infertility – a meta-analysis and literature review. *Reproductive Biomedicine Online* **15** 643–658.

Tüttelmann F, Simoni M, Kliesch S, Ledig S, Dworniczak B, Wieacker P & Röpke A 2011 Copy number variants in patients with severe oligozoospermia and Sertoli-cell-only syndrome. *PLoS ONE* **6** e19426. (doi:10.1371/journal.pone.0019426)

Tyler-Smith C & Krausz C 2009 The will-o'-the-wisp of genetics-hunting for the azoospermia factor gene. *New England Journal of Medicine* **26** 925–927. (doi:10.1056/NEJMe0900301)

Visscher PM, Brown MA, McCarthy MI & Yang J 2012 Five years of GWAS discovery. *American Journal of Human Genetics* **90** 7–24. (doi:10.1016/j.ajhg.2011.11.029)

Visser L, Westerveld GH, Korver CM, van Daalen SKM, Hovingh SE, Rozen S, van der Veen F & Repping S 2009 Y chromosome gr/gr deletions are a risk factor for low semen quality. *Human Reproduction* **24** 2667–2673. (doi:10.1093/humrep/dep243)

Vogt PH, Edelmann A, Kirsch S, Henegariu O, Hirschmann P, Kiesewetter F, Köhn FM, Schill WB, Farah S, Ramos C *et al.* 1996 Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Human Molecular Genetics* **5** 933–943. (doi:10.1093/hmg/5.7.933)

Wang XL, Mahaney MC, Sim AS, Wang J, Wang J, Blangero J, Almasy L, Badenhop RB & Wilcken DEL 1997 Genetic contribution of the

endothelial constitutive nitric oxide synthase gene to plasma nitric oxide levels. *Arteriosclerosis, Thrombosis, and Vascular Biology* **17** 3147–3153. (doi:10.1161/01.ATV.17.11.3147)

Wei B, Xu Z, Ruan J, Zhu M, Jin K, Zhou D, Xu Z, Hu Q, Wang Q & Wang Z 2012 MTHFR 677C>T and 1298A>C polymorphisms and male infertility risk: a meta-analysis. *Molecular Biology Reports* **39** 1997–2002. (doi:10.1007/s11033-011-0946-4)

Weiner AS, Boyarskikh UA, Voronina EN, Tupikin AE, Korolkova OV, Morozov IV & Filipenko ML 2014 Polymorphisms in folate-metabolizing genes and risk of idiopathic male infertility: a study on a Russian population and a meta-analysis. *Fertility and Sterility* **101** 87–94. (doi:10.1016/j.fertnstert.2013.09.014)

Wu W, Shen O, Qin Y, Lu J, Niu X, Zhou Z, Lu C, Xia Y, Wang S & Wang X 2012 Methylenetetrahydrofolate reductase C677T polymorphism and the risk of male infertility: a meta-analysis. *International Journal of Andrology* **35** 18–24. (doi:10.1111/j.1365-2605.2011.01147.x)

Xu K, Lu T, Zhou H, Bai L & Xiang Y 2010 The role of MSH5 C85T and MLH3 C2531T polymorphisms in the risk of male infertility with azoospermia or severe oligozoospermia. *Clinica Chimica Acta; International Journal of Clinical Chemistry* **411** 49–52. (doi:10.1016/j.cca.2009.09.038)

Xu M, Qin Y, Qu J, Lu C, Wang Y, Wu W, Song L, Wang S, Chen F, Chen H *et al.* 2013 Evaluation of five candidate genes from GWAS for association with oligozoospermia in a Han Chinese population. *PLoS ONE* **8** e80374. (doi:10.1371/journal.pone.0080374)

Yan L, Guo W, Wu S, Liu J, Zhang S, Shi L, Ji G & Gu A 2014 Genetic variants in nitric oxide synthase genes and the risk of male infertility in a Chinese population: a case–control study. *PLoS ONE* **9** e115190. (doi:10.1371/journal.pone.0115190)

Yatsenko AN, Roy A, Chen R, Ma L, Murthy LJ, Yan W, Lamb DJ & Matzuk MM 2006 Non-invasive genetic diagnosis of male infertility using spermatozoal RNA: KLHL10 mutations in oligozoospermic patients impair homodimerization. *Human Molecular Genetics* **15** 3411–3419. (doi:10.1093/hmg/ddl417)

Yatsenko AN, Georgiadis AP, Röpke A, Berman AJ, Jaffe T, Olszewska M, Westernströer B, Sanfilippo J, Kurpisz M, Rajkovic A *et al.* 2015 X-Linked TEX11 ,mutations, meiotic arrest, and azoospermia in infertile men. *New England Journal of Medicine* **372** 2097–2107. (doi:10.1056/NEJMoa1406192)

Ye JJ, Ma L, Yang LJ, Wang JH, Wang YL, Guo H, Gong N, Nie WH & Zhao SH 2013 Partial AZFc duplications not deletions are associated with male infertility in the Yi population of Yunnan Province, China. *Journal of Zhejiang University. Science. B* **14** 807–815. (doi:10.1631/jzus.B1200301)

Yen PH 2004 Putative biological functions of the DAZ family. *International Journal of Andrology* **27** 125–129. (doi:10.1111/j.1365-2605.2004.00469.x)

Yen PH, Chai NN & Salido EC 1996 The human autosomal gene DAZLA?: testis specificity and a candidate for male infertility *Human Molecular Genetics* **5** 2013–2017. (doi:10.1093/hmg/5.12.2013)

Yi W, Wu X, Lee T-H, Doggett NA & Her C 2005 Two variants of MutS homolog hMSH5: prevalence in humans and effects on protein interaction. *Biochemical and Biophysical Research Communications* **332** 524–532. (doi:10.1016/j.bbrc.2005.04.154)

Ying H, Scott MB & Zhou-Cun A 2012 Relationship of SNP of H2BFWT gene to male infertility in a Chinese population with idiopathic spermatogenesis impairment. *Biomarkers : Biochemical Indicators of Exposure, Response, and Susceptibility to Chemicals* **17** 402–406. (doi:10.3109/1354750X.2012.677066)

Yun YJ, Park JH, Song SH & Lee S 2008 The association of 4a4b polymorphism of endothelial nitric oxide synthase (eNOS) gene with the sperm morphology in Korean infertile men. *Fertility and Sterility* **90** 1126–1131. (doi:10.1016/j.fertnstert.2007.07.1382)

Zhang F, Li Z, Wen B, Jiang J, Shao M, Zhao Y, He Y, Song X, Qian J, Lu D *et al.* 2007 A frequent partial AZFc deletion does not render an increased risk of spermatogenic impairment in East Asians. *Annals of Human Genetics* **70** 304–313. (doi:10.1111/j.1529-8817.2005.00231.x)

Zhao H, Xu J, Zhang H, Sun J, Sun Y, Wang Z, Liu J, Ding Q, Lu S, Shi R *et al.* 2012 A genome-wide association study reveals that variants within the HLA region are associated with risk for nonobstructive azoospermia. *American Journal of Human Genetics* **90** 900–906. (doi:10.1016/j.ajhg.2012.04.001)

Zini A, Bryan MKO, Magid MS & Schlegel PN 1996 Immunohistochemical localization of endothelial nitric oxide synthase in human testis, epididymis, and Vas Deferens suggests a possible role for nitric oxide in spermatogenesis, sperm maturation, and programmed cell death. *Biology of Reproduction* **941** 935–941. (doi:10.1095/biolreprod55.5.935)

Zou S, Li Z, Wang Y, Chen T, Song P, Chen J, He X, Xu P, Liang M, Luo K *et al.* 2014 Association study between polymorphisms of PRMT6, PEX10, SOX5, and nonobstructive azoospermia in the Han Chinese population. *Biology of Reproduction* **90** 96. (doi:10.1095/biolreprod.113.116541)

Received 9 June 2015
First decision 9 July 2015
Revised manuscript received 17 July 2015
Accepted 27 July 2015