Published in final edited form as:

Nat Genet. 2009 July ; 41(7): 765–766. doi:10.1038/ng0709-765.

High Marks for GWAS

S Chanock

Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD, USA.

Abstract

Two genome-wide association studies have identified single nucleotide polymorphism (SNP) markers associated with the risk for testicular cancer. Both studies point towards variants in the vicinity of genes that have been implicated in testicular development, on chromosome 12, the ligand for the receptor tyrosine kinase, (*KITLG*) and on chromosome 5, sprouty 4 (*SPRY4*).

Every investigator conducting a genome-wide association study (GWAS) has a moment of pause before looking at their newly minted results. After all the effort to collect phenotypes and biospecimens, generate genotypes, and conduct association analyses, there is a moment of truth when one wonders whether it has been worth it. For Kanetsky et al and Rapley et al their hard work has certainly paid off. Their studies of testicular cancer reported in this issue have identified common single nucleotide polymorphism (SNP) markers on chromosome 5 and chromosome 12 associated with risk for testicular germ cell tumors (TGCT). Indeed, these two studies have put the genetics of TGCT on firmer ground with three new regions (the third on chromosome 6 was observed only by Rapley et al).

Some might argue that the genetics of TGCT was always on firm ground, it just took GWAS to conclusively identify markers for specific regions. A number of lines of evidence have pointed towards a strong hereditary component for TGCT. It is noteworthy that the two GWAS have found novel regions and that the estimated per allele odds ratio for the markers on chromosome 12 are the highest of any observed for a cancer phenotype, well above 2.0. While more work lies ahead to map, investigate and eventually develop the clinical utility of these genetic markers, the other good news here is that the two new hits, common to both studies, point towards regions harboring excellent candidate genes for study. Often, the ‘GWASer’ is faced with the perplexing challenge of what to make of the hits, but not here.

Finding Plausible Genes

In retrospect, it is not surprising that the locus on chromosome 12 was easily found. To detect a per allele odds ratio of 2.5 or greater, both studies were adequately powered. In the smaller study, Kanetsky et al scanned only 277 cases and compared them to 919 available controls. The Rapley study was slightly better powered because it scanned 730 cases and 1435 controls; the discovery set was with family-loaded cases. The SNP maps to a region harboring the *KITLG* gene (known as either Stem Cell or Steel Factor), which encodes the

ligand for the membrane receptor tyrosine kinase, c-KIT, a gene extensively investigated in multiple cancers. It is also a key actor in primordial germ cell development. Indeed, abundant corroborative data link somatic and germ-cell mutations to testicular cancer and infertility in mice and men, but no single piece accounts for the observed association. Even with such a promising candidate gene, one must invoke a healthy skepticism and consider the possibility that the risk alleles could work through a mechanism unrelated to this gene. Extensive follow-up is needed to provide the next level of plausibility for the observed association.

For the chromosome 5 finding, has chance favored two sets of prepared minds? In truth, they were lucky. The probability of both studies finding the same locus with an estimated per allele odds ratio of 1.3 is small because it is likely the product of the power of each relatively small study. Interestingly, in the recently reported GWAS of thyroid cancer, another disease with strong evidence for a significant heritable component, two regions were detected on chromosome 9q22.3 and 14q13.3, each with a prior candidate thyroid-specific gene based on a scan of only 192 cases but with 37,196 Icelandic ‘controls’, namely those without thyroid cancer. It is not surprising that the testicular and thyroid scans should find common variants given the strong familial component for each, but then again, the results are surprising given the small sample sizes and the observed odds ratios.

This is just the beginning.....

These findings are quite encouraging and should stimulate further primary scans as well as meta-analyses in TGCT. Additional susceptibility markers will likely emerge because GWAS are scalable, but probably with per allele odds ratios in the range observed for most cancer GWAS so far, namely 1.15 to 1.3. It would be surprising to find more high marks with odds ratios above 2.0, an exception to the more than 60 hits reported in nearly a dozen cancer GWAS. Other risk factors associated with testicular cancer should be explored to refine our understanding based on exposures and prior conditions (e.g., cryptorchidism, inguinal hernia and testicular atrophy) that are known to increase risk. The opportunity to explore early exposures, either prepubertal or in utero, and gene-environment interactions is rich, primarily because of the young age of onset of TGCT.

Once an association signal crosses the threshold of genome-wide significance, the next steps forward are often long and arduous. The mapping of each novel region with denser SNPs and in populations with differences in patterns of linkage disequilibrium, if available, should highlight variants for laboratory analyses using model systems, tissue samples, and in vitro techniques. The latter can provide plausibility for why certain variants contribute to the risk for TGCT, but rarely are sufficient to offer ‘proof’. It is more likely that the complexity of the genomic architecture could mirror the complexity of the disease. In other words, common, uncommon and rare variants could be lurking in these regions.

One of the daunting challenges will be to dissect the signals by molecular or histologic subtypes, as has been reported for estrogen receptor status in breast cancer. It is notable that the two scans combined seminoma and non-seminoma tumors, primarily to achieve adequate power to detect a main effect. Indeed, the findings indicate that both regions contribute to the

risk for TGCT overall, but the sub-analyses, while underpowered, were unable to distinguish a difference by subtype. It will take substantially larger data sets, probably through combined analyses with all available studies to ferret out the contribution of genetics to different sub-types.

Recently, genome-wide association studies have come under criticism for failing to adequately explain the heredity underlying a spectrum of human diseases and traits. This is somewhat surprising in light of the remarkable discovery of hundreds of new loci for human diseases and traits, all using the GWAS approach. We are still early in the discovery phase and it is daunting to assess the actual contribution of genetic variation to disease until we know more about the spectrum of the contributing variants and the environmental risk factors. Furthermore, we need to be cautious about quickly transitioning the newly discovered loci to use in tools for clinical risk assessment, often in the same studies, which may not necessarily be well designed to address the question of risk. Indeed, we have a long way to go to make sense of the high marks garnered by GWAS this far.

Proposed References:

1. Kanetsky. et al. Nature Genetics. 2009
2. Rapley et al., Nature Genetics 2009
3. Henderson BE, et al. In Urological Cancer (ed Skinner DG) (pp237–250) 1983
4. Gudmundsson et al. Nature Genetics 41: 460–464, 2007.
5. Easton DF and Eeles RA. Hum Mol Genet 17:R109–115, 2008. [PubMed: 18852198]
6. Garcia-Closas M et al. Plos Genetics 4:e1000054, 2008. [PubMed: 18437204]
7. Manolio et al. J Clin Invest 118:1590–1605, 2008. [PubMed: 18451988]
8. Kraft et al. Nature Reviews Genetics 10:264–269, 2009.