Genome-wide association studies and musculoskeletal diseases

Patrick Danoy & Matthew A Brown†

†Author for correspondence: Diamantina Institute of Cancer, Immunology and Metabolic Medicine, The University of Queensland, Princess Alexandra Hospital, Ipswich Road, Woolloongabba, Qld, 4102, Australia ■ Tel.: +61 732 402 870 ■ Fax: +61 732 405 946 ■ matt.brown@uq.edu.au

Bone and joint diseases are major causes of morbidity and mortality worldwide, and their prevalence is increasing as the average population age increases. Most common musculoskeletal diseases show significant heritability, and few have treatments that prevent disease or can induce true treatment-free, disease-free remission. Furthermore, despite valiant efforts of hypothesis-driven research, our understanding of the etiopathogenesis of these conditions is, with few exceptions, at best moderate. Therefore, there has been a long-standing interest in genetics research in musculoskeletal disease as a hypothesis-free method for investigating disease etiopathogenesis. Important contributions have been made through the identification of monogenic causes of disease, but the holy grail of human genetics research has been the identification of the genes responsible for common diseases. The development of genome-wide association (GWA) studies has revolutionized this field, and led to an explosion in the number of genes identified that are definitely involved in musculoskeletal disease pathogenesis. However, this approach will not identify all common disease genes, and although the current progress is exciting and proves the potential of this research discipline, other approaches will be required to identify many of the types of genetic variation likely to be involved.

The genome-wide association (GWA) study era really only began in June 2007, with the publication of the findings of the Wellcome Trust Case Control Consortium (WTCCC). The WTCCC showed that it is possible to identify susceptibility alleles and replicate these findings even with minor effects on disease risk in a systematic, unbiased, hypothesis-free approach [1,2]. Whilst the WTCCC papers were not the first GWA study to be reported, they made a critical contribution by setting the standard for the performance of these studies, addressing and discussing issues such as quality control, design and analysis.

Two recent, but very important, technological and scientific advances made GWA studies possible. The first was the completion of the International HapMap project that describes and annotates the common patterns of DNA sequence variation and linkage disequilibrium (LD) structure in the human genome [3]. This resource has made the intelligent design of marker panels covering the genome as efficiently as possible feasible through the use of tag single nucleotide polymorphisms (tagSNPs) in regions of high LD. By selecting SNPs that tag regions of the genome, more than 80% of the genome can be screened, whilst only typing approximately 10% of the common SNPs present. The second critical advance was the development of high-throughput microarray SNP genotyping platforms, which allow the genotyping of hundreds of thousands (up to more than a million) of polymorphisms across the genome in large cohorts at a low cost per SNP and with a very low error rate. Nonetheless, GWA studies, while now financially feasible, still remain very expensive experiments and are usually the result of major international collaborations.

The WTCCC undertook two separate experiments. In the first, they genotyped 14,000 cases of seven common diseases (including rheumatoid arthritis [RA]) and 3000 shared controls using a panel of 500,000 evenly scattered SNPs throughout the genome. In the second experiment, they set out to type 14,500 nonsynonymous SNPs (nsSNPs) in 6000 unrelated affected individuals of four diseases (including ankylosing spondylitis [AS]) and 1500 common controls. There are lessons to be learnt about this type of study in both cases. In RA, these experiments have strongly replicated previously identified susceptibility loci, namely the *HLA-DRB1* and *PTPN22* loci. They also identified nine SNPs with high association score statistics, but not reaching genome-wide significance ($1 \times 10^{-5} < p < 1 \times 10^{-7}$). Of particular interest amongst these variants are potential associations with subunits of the IL2 receptor (*IL2RA* and *IL2RB*), and the sex-differentiated association of rs11761231 on chromosome seven ($p = 6.8 \times 10^{-8}$) with an additive odds ratio for females of 1.32. For another, rs6920220 on 6q23,

Keywords

ankylosing spondylitis ■ autoimmunity ■ GWA studies ■ musculoskeletal ■ osteoporosis ■ rheumatoid arthritis ■ rheumatology

Special Report
Future Rheumatology

future medicine part of fsg

association has since been replicated [4], while a new variant in the region (less than 4 kb away) has also been identified in an independent study including a replication phase [5]. The identity of the primary associated gene is unknown, and may be either *TRAF1* or *C5*, or potentially both. Thus, although association mapping has higher resolution than linkage mapping, it is still a nontrivial challenge to achieve certainty, using genetic association studies alone, of the identity of the primary associated variant, and in some cases, even of the primary associated gene. A meta-analysis of these studies and comprehensive replication study is eagerly awaited.

Using the nsSNPs genotyping approach, the long and well-documented major histocompatibility complex (MHC) association with AS was easily confirmed. An additional two loci, *ERAP1* (*ARTS1*) and *IL23R*, were also found to be new susceptibility regions for disease. Both loci showed strong association in the original screen population, replication and combined analyses. The products of these two genes represent excellent biological candidates for association with AS. The protein ERAP1 is involved in trimming peptides for MHC class I presentation [6], but also cleaves cell surface receptors for pro-inflammatory cytokines [7]. IL23R is a key factor in the regulation of the Th17 cells that express high levels of proinflammatory cytokines [8], and animal studies have shown that blocking IL-23 reduces inflammation [9]. Genetic variants that vary the function of both ERAP1 and IL23R could very well have a major effect on the inflammatory pathogenesis of AS; neither had been considered by the immunological community as being of significance in the disease prior to the WTCCC findings. Indeed, the *IL23R* finding has been received with skepticism in some quarters, as immunological dogma has it that AS is a human leukocyte antigen (HLA) class I restricted disease, and that therefore only mechanisms involving CD8 T cells could have a major role. Clearly, it is time to change that belief.

Another feature of these studies has been the breadth of association observed between the immunological diseases studied and the MHC. Significant association ($p < 10^{-7}$) was observed for over 500 kb around the MHC in RA, AS, and in other immunologically mediated diseases studied (Type 1 diabetes, autoimmune thyroid disease and multiple sclerosis). The breadth of association suggests either that LD in this area is extremely long-ranging, that there is more than one disease-associated MHC gene in each of these conditions, or most likely, a combination of both explanations. There is already considerable evidence for the existence of non-*HLA-DRB1* MHC associations in RA [10–12], and non-*HLA-B27* MHC associations in AS [13,14]. The recent demonstration, using a combination of classical HLA and dense SNP genotyping, that *HLA-A* and *HLA-B* are independently associated with Type 1 diabetes mellitus [15], a disease hitherto thought to be HLA class II-restricted, suggests that further such studies in other MHC-associated immunologically-mediated diseases would be very worthwhile.

More recent work by others on the quantitative traits that influence risk for osteoporosis has identified several loci for which there is strong evidence of association with bone mineral density and fracture [16,17]. Four of these were already known as good candidate genes (*ESR1*, *LRP5*, *OPG* and *RANKL*), and the other two are new loci at 1p36 (near ZBTB40) and 6p21 in the MHC. While these variants have low odds ratios, four of them lie within crucial regulators of bone homeostasis, and the results thus provide insight into the underlying biology of disease. Furthermore, the combination of SNPs in *LRP5* and *OPG* was associated with an odds ratio for fracture of 1.33 and was common (present in 22% of the cohort), indicating a significant population effect size. Prior to these studies, it had been thought that the low genetic covariance between fracture and bone density indicated that few genes would influence both traits [18,19]. Contrary to these predictions, no genes have yet been identified that affect fracture risk but not bone density, whereas most but not all genes associated with bone density affect fracture risk. This is not surprising as, with the exception of younger patients, fracture risk has much lower heritability than bone density [18–23]. Even these early and underpowered studies have thus made significant contributions, identifying regions not previously known to be involved in bone fragility, and identifying SNP combinations likely to be of value in fracture risk prediction.

Therefore, in a very short period, GWA studies have become established as the method of choice for investigation of the genetic architecture of common diseases. GWA studies are now underway in most common musculoskeletal diseases, and have thankfully largely ended the phase of underpowered candidate gene studies that dominated the field for much of the past decade. These studies added little to our understanding of the etiopathogenesis of disease, being generally too small to provide robust findings,

either positive or negative. It is to be hoped and encouraged that GWA studies do not repeat some of the errors of this era, such as 'salami slicing' of data (where data gathered by one study are separately reported in multiple end publications), and overinterpretation of modest findings. Given the cost of these studies and the large contribution we ask of the cohorts that are required to be studied, the authors feel it is an ethical requirement of investigators to design optimal studies with adequate power, preplanned confirmation studies and early public availability of genotype data to *bona fide* researchers to maximize its utility. The two WTCCC studies demonstrate that a sample size of 2000 cases and 3000 controls is about the smallest sample size for a single phase case–control design that has adequate power to detect the likely genetic effects present in most common diseases. Smaller sample sizes, such as those used in the nsSNP component of the WTCCC (1000 cases and 1500 controls), were enough to pick up quite large genes (e.g., *ERAP1* and *IL23R*). However, the fact that in the other three diseases included in this study (autoimmune thyroid disease, multiple sclerosis and breast cancer) no non-MHC genes achieved genome-wide significance indicates that this sample size is too small to investigate most common diseases (Figure 1).

One frequent criticism of these studies is that the genes that have been identified have had only been of small magnitude. Of course, this is not unexpected, as selection pressure makes it unlikely that genes of major adverse effect can become common in the general population. Nonetheless, it also raises several points of common confusion that would be valuable to clarify. Firstly, it is important to note that the current studies generally have not identified the key associated variants, but are only tagSNP studies. The true disease-associated variants, once identified, may have stronger levels of association than the tagSNPs used in LD mapping. For example, in the WTCCC nsSNP study of AS, the strongest MHC-associated SNPs had odds ratios of only 3–3.5. We know that the main MHC allele in AS is *HLA-B27*, which in white European populations has an odds ratio for disease of more than 100. Secondly, the odds ratio is a poor measure of the overall contribution of a genetic variant to the overall risk of disease. This contribution depends on the prevalence of the associated allele; common disease-associated alleles have greater effects in the population. The contribution of a genetic variance for dichotomous traits is better reported using the population-attributable risk fraction (the proportion of cases of a disease that would not occur if the effect of that variant was removed from the population) (Figure 2). For quantitative traits, the proportion of the variance of that trait due to that specific genetic variant can be reported. Lastly, the contribution of a gene to a disease depends not only on the importance of a gene in the etiopathogenesis of a disease, but also on the amount of functionally relevant genetic variation in that gene. Thus, genes like *TNF* may be essential to the development of inflammation in autoimmunity, but because there is little or no effect of genetic variants of *TNF* on TNF expression or function, they have little or no association with disease on a population level.

The WTCCC experiments also provided researchers with insights into population structure, sample sizes required to detect moderate genetic effects and methodological improvements. The study implemented extensive quality control checks in DNA quality and allele calling (Q–Q plots to identify inflation of association findings, thresholds for SNP, individual genotyping success rates, and, for Hardy–Weinberg equilibrium tests, visual inspection of cluster plots of associated variants) in order to minimize incorrect genotype calls and false-positive signals. Another issue addressed by the study was the use of a common set of controls for a number of different diseases. What would be

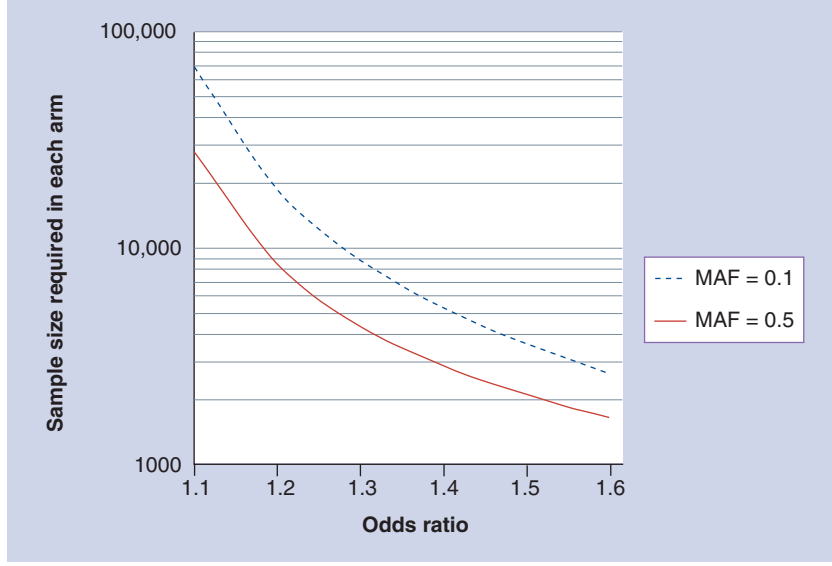

Figure 1. Sample size requirements to identify genetic variants with varying odds ratio. Sample size requirements to identify genetic variants with varying odds ratios, assuming average D′ = 0.8, disease population prevalence = 1%, power = 80%, and significance threshold $p = 10^{-7}$, an additive genetic model, and an equal number of cases and controls. For the genetic effect, sizes likely to be operating in common diseases, more than 2000 cases and controls are required to achieve adequate study power (calculated using 'Genetic Power Calculator' [33]). MAF: Minor allele frequency.

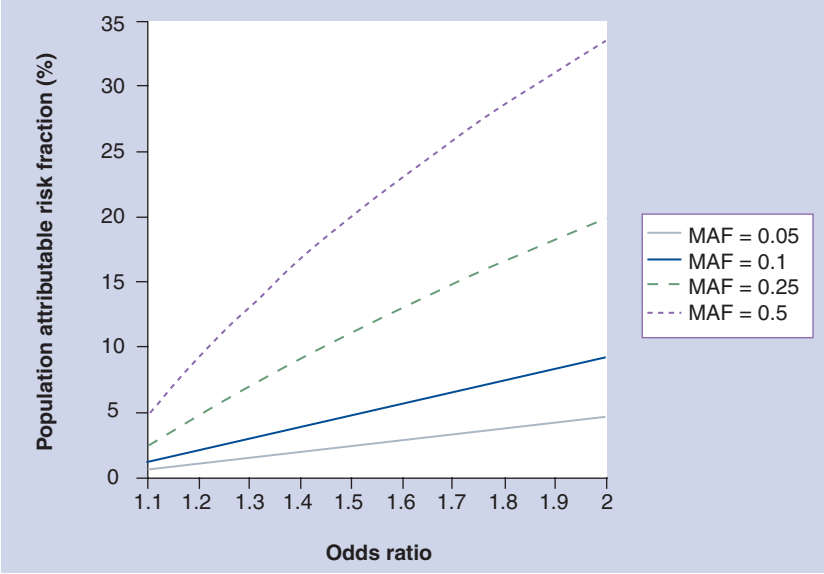

Figure 2. Population-attributable risk fraction given odds ratio and minor allele frequency. The contribution of a genetic variant to the risk of a disease is related both to the odds ratio and to its frequency.
MAF: Minor allele frequency.

the consequences of failing to match cases and controls for socio-demographic variables, or the impact of misclassification bias when controls meet the criteria used to define cases? The study demonstrated that both concerns were overstated and had little effect on power. Using these common controls, the study identified 13 loci that vary in minor allele frequency based on geographical regions across the UK, but concluded that population structure has, at the most, a minor effect, providing that non-European individuals were excluded from analysis. Major cost-cutting benefits can be expected from the use of common controls. Databases with more than 5000 white European controls are already available, and the WTCCC will shortly add a further 6000 controls to the public resource, typed at around 1 million SNPs each. If working in ethnically matched populations, unless one is interested in rare variants, or in a trait that is extremely common in the general community (≥20%), the available historic controls are adequate in most cases.

Whilst the initial WTCCC GWA study did not include a confirmation phase, it has now become standard to test findings for replication in a second cohort. Cost-effective phased designs have been developed that employ smaller discovery cohorts, but follow-up a large number of SNPs from that phase in a larger and more powerful replication set. These replication findings are then generally analyzed both independently with the discovery set, and as a combined analysis. This approach significantly reduces the amount of genotyping involved. The economics of the design depend on the relative costs of the fixed marker genome-wide discovery chips, and the custom-designed confirmation chips. As the cost of fixed marker chips has declined faster than custom genotyping, the value of phased designs has also declined. Further reduction in genotyping cost can be achieved by the use of common or historic controls, and evidence has also been provided that sample pooling can be robust and sensitive enough to identify risk variants [24,25]. Disadvantages of DNA pooling include reduced genotyping accuracy (which tends to be over-represented amongst apparently associated SNPs), the absence of individual level data precluding haplotyping and complicating imputation of nongenotyped variants, and increased challenges in typing copy number variants.

GWA studies are not the solution to all gene-mapping challenges. In particular, they clearly struggle to identify susceptibility loci belonging to highly heterogeneous traits, or where the tagSNP approach has low power, such as where the common variant/common disease hypothesis breaks down, in regions of low LD, and for copy number variants. It is obvious from inspection of the number of genes identified in different diseases in the WTCCC study that diseases that were *a priori* likely to be more complex, such as bipolar disorder and hypertension, had fewer genes identified than relatively phenotypically homogenous diseases, such as Crohn's disease and Type 1 diabetes. These heterogenous diseases may be tractable with larger sample sizes, but a more efficient approach will be to try to study tighter clinical subsets to minimize heterogeneity. This has relevance to musculoskeletal diseases such as osteoarthritis and osteoporosis, which are likely to be genetically very heterogeneous. Whilst the WTCCC demonstrated that for many genes the common disease–common variant hypothesis is likely to be true, it is not clear what proportion of genes this holds true for. It remains possible that for a significant proportion of genes multiple variants are involved in the population risk of a disease, and these will be difficult, if not impossible, to detect by GWA approaches. Instead, resequencing approaches will be needed to identify rare alleles in large cohorts. There is also a great deal of interest in the genetics community in other types of genetic variation, such as copy number variant and methylation changes, which are at best poorly typed by tagSNP approaches. High-throughput methods for studying these genetic variants are still in development, and

it is likely that we will see a further explosion of findings once these genotyping and analysis methods reach maturity. Lastly, current GWA studies have largely ignored gene–gene interaction and gene–environment interaction. Both are likely to contribute significantly to the risk of developing disease, and more powerful studies and better statistical approaches will be needed to investigate them properly.

Amongst the more interesting findings to come out of GWA studies has been the overlap between diseases in their underlying genetic architecture. This has been particularly apparent for autoimmune diseases, where several genes with pleiotropic effects across different diseases have been discovered. Researchers are thus increasingly investigating associations from related diseases to identify disease-associated variants in other diseases. Examples of this include the association of variants of *PTPN22* with systemic lupus erythematosus, RA, Type 1 diabetes and autoimmune thyroid disease [26–29], *STAT4* in systemic lupus erythematosus and RA [30], and association of *IL23R* with Crohn's disease, psoriasis and AS [2,31,32]. It therefore seems that certain genes may act as master regulators of the autoimmune system, and detailed studies of the newly identified variants across multiple autoimmune diseases may help explain both the commonalities and differences among these diseases. It is also notable that some of these variants were not picked up in fixed-marker GWA studies, as they lay in areas poorly tagged by the available chips. Thus, there remains a role for candidate gene studies, particularly in those areas with poor coverage in GWA studies.

The identification of risk alleles will go a long way towards elucidating the biological processes driving musculoskeletal diseases. These investigations are still in their early stages, and have as yet identified only a small proportion of the determinants of the genetic risk. Nonetheless, even in these early days of the GWA era, it is clear that these hypothesis-free approaches are a powerful method to investigate disease pathogenesis, and that the longstanding promise of human genetics is now being delivered. The next challenge lies with functional biologists to determine the mechanism by which the associated variants influence disease risk, and to develop interventions based on that information. This will also be a nontrivial exercise, but at least the hypothesis-driven research community will now have a firm foundation on which to build.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Executive summary

- The International HapMap project (to select tagSNPs) and availability of high-throughput genotyping platforms have made genome-wide association (GWA) studies technically and financially possible.
- The Wellcome Trust Case Control Consortium studies have addressed and discussed many issues (quality control, design and analysis) and have paved the way forward for properly powered GWA studies.
- Population stratification through the use of common controls has little effect on study power if ethnicity-matched cases and controls are used.
- GWA studies may only identify genes or broader genetic regions and further fine mapping/resequencing analysis and functional analysis will be needed to identify causative genetic variants.
- Rhematoid arthritis and osteoporosis GWA studies have identified potential and confirmed associations with good candidate genes and provide insights into disease etiopathogenesis.
- The novel *ERAP1* and *IL23R* findings should change the dogma that ankylosing spondylitis is solely a human leukocyte antigen class I mediated disease and implicate a role for Th17 cells in disease.
- Genes with pleiotropic effects across different diseases suggest an overlap in underlying genetic architecture of autoimmune diseases.

Bibliography

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1. Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447(7145), 661–678 (2007).

▪▪ **Breakthrough paper that brought genome-wide association studies to maturity.**

2. Burton PR, Clayton DG, Cardon LR *et al.*: Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat. Genet.* 39(11), 1329–1337 (2007).

3. International HapMap Consortium: A haplotype map of the human genome. *Nature* 437(7063), 1299–1320 (2005).

▪▪ **A landmark paper in human genetics that underpins much of our strategies in common disease gene mapping today.**

4. Thomson W, Barton A, Ke X *et al.*: Rheumatoid arthritis association at 6q23.

Nat. Genet. 39(12), 1431–1433 (2007).

5. Plenge RM, Cotsapas C, Davies L *et al.*: Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nat. Genet.* 39(12), 1477–1482 (2007).
6. Saric T, Chang SC, Hattori A *et al.*: An IFN-γ-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat. Immunol.* 3(12), 1169–1176 (2002).
7. Cui X, Rouhani FN, Hawari F, Levine SJ: An aminopeptidase, ARTS-1, is required for interleukin-6 receptor shedding. *J. Biol. Chem.* 278(31), 28677–28685 (2003).
8. Chen Z, O'Shea JJ: Regulation of IL-17 production in human lymphocytes. *Cytokine* 41(2), 71–78 (2008).
9. Cua DJ, Sherlock J, Chen Y *et al.*: Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. *Nature* 421(6924), 744–748 (2003).
10. Harney SMJ, Vilarino-Guell C, Adamopoulos I *et al.*: Fine mapping of the MHC class III region demonstrates association of AIF1 and rheumatoid arthritis. *Rheumatology (Oxford)* (2008) (Epub ahead of print).
11. Jawaheer D, Li W, Graham RR *et al.*: Dissecting the genetic complexity of the association between human leukocyte antigens and rheumatoid arthritis. *Am. J. Hum. Genet.* 71(3), 585–594 (2002).
12. Newton JL, Harney SM, Timms AE *et al.*: Dissection of class III major histocompatibility complex haplotypes associated with rheumatoid arthritis. *Arthritis Rheum.* 50(7), 2122–2129 (2004).
13. Brown MA, Kennedy LG, Darke C *et al.*: The effect of *HLA-DR* genes on susceptibility to and severity of ankylosing spondylitis. *Arthritis Rheum.* 41(3), 460–465 (1998).
14. Sims AM, Barnardo M, Herzberg I *et al.*: Non-B27 MHC associations of ankylosing spondylitis. *Genes Immun.* 8(2), 115–123 (2007).
15. Nejentsev S, Howson JM, Walker NM *et al.*: Localization of Type 1 diabetes susceptibility to the MHC class I genes *HLA-B* and *HLA-A*. *Nature* 450(7171), 887–892 (2007).

▪ **Shows how to map MHC genes in the presence of more than one susceptibility locus.**

16. Richards JB, Rivadeneira F, Inouye M *et al.*: Bone mineral density, osteoporosis and osteoporotic fractures: a genome-wide association study. *Lancet* 371(9623), 1505–1512 (2008).
17. Styrkarsdottir U, Halldorsson BV, Gretarsdottir S *et al.*: Multiple genetic loci for bone mineral density and fractures. *N. Engl. J. Med.* 358(22), 2355–2365 (2008).
18. Andrew T, Antioniades L, Scurrah KJ, Macgregor AJ, Spector TD: Risk of wrist fracture in women is heritable and is influenced by genes that are largely independent of those influencing BMD. *J. Bone Miner. Res.* 20(1), 67–74 (2005).
19. Deng HW, Mahaney MC, Williams JT *et al.*: Relevance of the genes for bone mass variation to susceptibility to osteoporotic fractures and its implications to gene search for complex human diseases. *Genet. Epidemiol.* 22(1), 12–25 (2002).
20. Kannus P, Palvanen M, Kaprio J, Parkkari J, Koskenvuo M: Genetic factors and osteoporotic fractures in elderly people: prospective 25 year follow up of a nationwide cohort of elderly Finnish twins. *BMJ.* 319(7221), 1334–1337 (1999).
21. MacGregor A, Snieder H, Spector TD: Genetic factors and osteoporotic fractures in elderly people. Twin data support genetic contribution to risk of fracture. *BMJ.* 320(7250), 1669–1670 (2000).
22. Michaelsson K, Melhus H, Ferm H, Ahlbom A, Pedersen NL: Genetic liability to fractures in the elderly. *Arch. Intern. Med.* 165(16), 1825–1830 (2005).
23. Deng HW, Chen WM, Recker S *et al.*: Genetic determination of Colles' fracture and differential bone mass in women with and without Colles' fracture. *J. Bone Miner. Res.* 15(7), 1243–1252 (2000).
24. Baum AE, Akula N, Cabanero M *et al.*: A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol. Psychiatry* 13(2), 197–207 (2008).
25. Steer S, Abkevich V, Gutin A *et al.*: Genomic DNA pooling for whole-genome association scans in complex disease: empirical demonstration of efficacy in rheumatoid arthritis. *Genes Immun.* 8(1), 57–68 (2007).
26. Begovich AB, Carlton VE, Honigberg LA *et al.*: A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (*PTPN22*) is associated with rheumatoid arthritis. *Am. J. Hum. Genet.* 75(2), 330–337 (2004).
27. Bottini N, Musumeci L, Alonso A *et al.*: A functional variant of lymphoid tyrosine phosphatase is associated with Type I diabetes. *Nat. Genet.* 36(4), 337–338 (2004).
28. Kyogoku C, Langefeld CD, Ortmann WA *et al.*: Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *Am. J. Hum. Genet.* 75(3), 504–507 (2004).
29. Velaga MR, Wilson V, Jennings CE *et al.*: The codon 620 tryptophan allele of the lymphoid tyrosine phosphatase (*LYP*) gene is a major determinant of Graves' disease. *J. Clin. Endocrinol. Metab.* 89(11), 5862–5865 (2004).
30. Remmers EF, Plenge RM, Lee AT *et al.*: *STAT4* and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N. Engl. J. Med.* 357(10), 977–986 (2007).
31. Cargill M, Schrodi SJ, Chang M *et al.*: A large-scale genetic association study confirms *IL12B* and leads to the identification of *IL23R* as psoriasis-risk genes. *Am. J. Hum. Genet.* 80(2), 273–290 (2007).
32. Duerr RH, Taylor KD, Brant SR *et al.*: A genome-wide association study identifies *IL23R* as an inflammatory bowel disease gene. *Science* 314(5804), 1461–1463 (2006).
33. Purcell S, Cherny SS, Sham PC: Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19(1), 149–150 (2003).

Affiliations

▪ Patrick Danoy
Diamantina Institute of Cancer, Immunology and Metabolic Medicine, The University of Queensland, Princess Alexandra Hospital, Ipswich Road, Woolloongabba, Qld, 4102, Australia

▪ Matthew A Brown
Diamantina Institute of Cancer, Immunology and Metabolic Medicine, The University of Queensland, Princess Alexandra Hospital, Ipswich Road, Woolloongabba, Qld, 4102, Australia
Tel.: +61 732 402 870;
Fax: +61 732 405 946;
matt.brown@uq.edu.au