Twin Research and Human Genetics (2020), **23**, 90–93
doi:10.1017/thg.2020.38

CAMBRIDGE UNIVERSITY PRESS

Article

Gene Discovery Using Twins

David Duffy[1], Rick Sturm[2], Gu Zhu[1] and Stuart MacGregor[3]

[1]Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia, [2]The University of Queensland Diamantina Institute, The University of Queensland, Brisbane, Australia and [3]Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia

Abstract

One of Nick's key early achievements at QIMR was to establish a twin study on melanoma risk factors. The Brisbane Twin Nevus Study (BTNS) had an initial focus on nevus (mole) count in adolescents but, reflecting Nick's broad interests, expanded in scope enormously over the decades. In the skin cancer arena, BTNS was essential to genetic discoveries in melanoma, eye color and pigmentation. Later studies amassed data on thousands of phenotypes, ranging from molecular phenotypes such as gene expression to studies where gene mapping findings in adolescents turned out to have translational potential in late-onset diseases. Nick's twin data have formed the basis for an enormous range of discoveries, with Nick and his colleagues continuing to capitalize on these data.

Keywords: Eye disease; genetics; genomewide association study; melanoma; risk factors; twin study

(Received 24 February 2020; accepted 25 March 2020)

When Nick first set up his laboratory at QIMR, it was inevitable that he would work on the genetics of melanoma in collaboration with Adele Green and Bob MacLennan. After all, Queensland regularly holds the honor of having the highest incidence of the disease in the world (swapping occasionally with New Zealand). The Queensland Melanoma Project (at Princess Alexandra Hospital) had previously published an estimate for the overall heritability of melanoma at 10%, based on local familial recurrence risks, and densely affected pedigrees around the world were being collected at that time to detect major risk genes via genetic linkage analysis. Adele (and Bob) had a track record in studies of the number of acquired melanocytic nevi (common moles) on the skin, a potent melanoma risk factor, looking at adults but also in children where these lesions first make their appearance and increase in number rapidly in adolescence. Nick was fully aware of the need for large sample sizes to obtain adequate statistical power for genetic studies and had come with the practical experience of (co-) founding the Australian Twin Registry (ATR). So it was natural, he was involved in the design and running of two big population-based studies. One was a pedigree-based study of melanoma — The Queensland Familial Melanoma Project (QFMP) — ascertaining all incident cases in the state in a two-year period, and aimed at segregation and linkage analysis. The other was a classical twin study of mole counts and other melanoma risk factors such as skin, hair and eye color using 12-year-old schoolchildren, who it was planned to follow up until at least age 16. The biometrical types of genetic analyses Nick was an expert in included modeling the genetics of multivariate and time series data — we were just moving to use of the structural equation modeling package LISREL for this kind of dataset. We should mention that Nick's previous omnivorous interests in genetics definitely included pigmentation genetics even though he would not have known much about melanoma.

The latter design of recruiting twins from schools is a classic way to find twins in the population and had been in use since the 1920s. It was not until later that we found out that one of the very first classical twin studies of any trait ever had been of mole counts, carried out by the dermatologist Siemens, and described in his 1924 book *Die ZwillingsPathologie* (the heritability was 40% or so). The 'Canberra' study that Nick had run (that became the core of the ATR) had followed the model of the big 1960s Scandinavian twin registries of collecting as many different phenotypes as possible using questionnaires. So the Brisbane Twin Nevus Study (BTNS) also included a wide range of psychological and health variables for which the genetics could be studied essentially 'for free'.

So the BTNS was in some respects fairly straightforward. Every year between 1992 and 2016 a fixed number of twins turned 12 in South-East Queensland within the participating schools, and on average 80 families would agree to take part. The parents would bring the twins in to QIMR for a visit — a well-known feature of being a twin or having twin children is an interest in genetics and participating in research. There, the study nurses would count all pigmented lesions >2 mm diameter on their skins (aside from obvious freckles!), measure skin reflectance and assess pigmentation. In passing, Nick and Adele had an ambitious project (McGregor et al., 1999) to digitally image all the moles on a subset of the BTNS, which anticipated modern dermatological tools such as the FotoFinder and Canfield Vectra systems. The twins and accompanying parent(s) would complete voluminous questionnaires and perform in various psychological tests and have blood collected for DNA extraction and a battery of biochemical and hematological assays. A smaller number of twins would return at ages 14 and 16 years. Later on, the study ramified as twins

Author for correspondence: Stuart MacGregor, Email: Stuart.MacGregor@qimrberghofer.edu.au

Cite this article: Duffy D, Sturm R, Zhu G, and MacGregor S. (2020) Gene Discovery Using Twins. *Twin Research and Human Genetics* **23**: 90–93, https://doi.org/10.1017/thg.2020.38

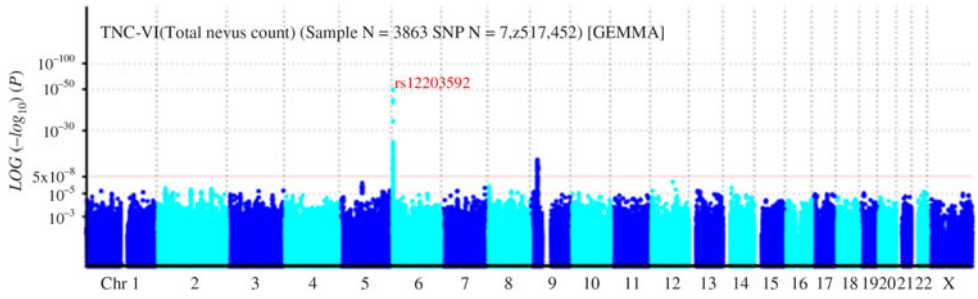

Fig. 1. GWAS for total nevus count in 12-year-old twins from Brisbane Twin Nevus Study as of January 2020.

themselves had children and funding became available for further follow-up. It also changed its name to reflect this, becoming the Brisbane Longitudinal Twin Study.

Although our main interest in collecting twins was to carry out classical twin analyses, we were also looking to the possibility of (dizygotic) sib-pair linkage analysis, again an approach requiring the samples sizes we were collecting. In our 1999 paper (Zhu et al., 1999), we reported on results from candidate microsatellite (and even a couple of single-nucleotide polymorphism [SNP]) marker genotyping of 352 families (a number that now seems hilariously small), where we found a quantitative trait locus close to the familial melanoma gene *CDKN2A* explained 27% of total variance in total body mole count (lod 2.6, $p = 6 \times 10^{-4}$). We concluded at the time that the actual nevus locus must be a common regulatory variant close to *CDKN2A*, probably centromeric to that gene, but despite follow-up fine mapping by linkage and then association analysis, did not successfully localize it until our UK collaborators pointed to a SNP in the *MTAP* gene, actually telomeric to *CDKN2A* — the joint paper (Falchi et al., 2009) describing this came out in 2009. In our more recent melanoma genomewide association study (GWAS) meta-analyses, we confirmed the association of both mole count and melanoma with these common *MTAP* alleles. Mechanistically, it is still not completely clear how these act — there is evidence implicating *MTAP* itself as important in carcinogenesis, even though *CDKN2A* is such a good candidate. In the same paper, we also partook in the discovery of the *PLA2G6* as another locus for mole count and melanoma risk — interestingly this has little effect on mole count in 12-year olds, but is quite easy to detect in adult samples. Its effects on melanoma risk were confirmed in the extended QIMR melanoma panel of studies that was based on the QFMP.

One of the other melanoma risk phenotypes we studied in the BTNS was eye color, with the advent of digital photography in the early 2000s greatly enhancing the phenotypic characteristics of the iris that could be captured and studied in the twins. It had been known that blue eye color was under the control of a high penetrance recessive locus, but the actual gene had not been positively identified, though a linkage analysis in 1996 had pointed to the vicinity of the oculocutaneous albinism 2 gene (*OCA2*) on chromosome 15. Using the twins, we first published a combined segregation-linkage analysis of 525 BTNS families (Zhu et al., 2004), confirming linkage to a microsatellite marker close to *OCA2*, and showing that this locus explained 75% of the population variance in eye color. By 2007 (Duffy et al., 2007), we had identified a three-SNP haplotype that almost completely explained blue eye color that lay not in *OCA2* itself but in an intron of the neighboring *HERC2* gene, and in 2008 published simultaneously with two other groups that rs12913832 was the single causative locus (Sturm et al., 2008).

As noted above, we also had an interest in several other pigmentation loci. Rick Sturm, then at University of Queensland, was one of the eminences in the study of human pigmentation genetics and the melanocortin-1 receptor (gene *MC1R*) in particular. Valverde and colleagues (1995) had reported variants in *MC1R* were associated with fair skin and red hair and showed one variant also predicted melanoma risk (Valverde et al., 1995, 1996). Rick and Nick's first paper together (Box et al., 1997) also came out in 1997 (Rick vividly remembers sitting at the computer with Nick as the latter did all the analyses), where they reported results of sequencing *MC1R* in red-headed twins from the BTNS, discovering a number of new variant alleles. This work was soon followed up by studies using both the BTNS but also the QFMP. Indeed, the BTNS families were often used as controls in melanoma case–control analyses, meaning the statistical methods had to incorporate the relatedness of the control samples. In Palmer et al. (2000), we showed that risk of melanoma due to carrying *MC1R* red-hair variants was not completely explainable by measured skin color. This finding has been extensively replicated, and it was clear that the effects of *MC1R* are not just on coloring (changing the melanin in the skin from dark eumelanin to a mixture of eumelanin and reddish-brown pheomelanin), but also the cell cycle and DNA repair in the melanocyte. In another paper (Box et al., 2001), we demonstrated that *MC1R* genotype modifies the penetrance of high-risk *CDKN2A* mutations, a finding that we thought straightforward, coming from the study of polygenic traits in twins, but the magnitude of this effect is of great interest to clinicians. We also classified the large number of *MC1R* coding variants into high-red-hair penetrance alleles (*R* alleles) and low penetrance *(r* alleles) using our BTNS data, with *R*/*R* being very likely to be redheads (Sturm et al., 2003).

Returning to the BTNS, we looked for effects of variants in known pigmentation genes on mole count. *MC1R* did not seem to have much effect, though in more recent analyses we see that counts increase in compound heterozygotes (e.g., *R/r)*, but fall in *R*/*R* homozygotes. In the case of the rs12203592 polymorphism in *IRF4*, associated with skin and hair color by us and others in a multicountry consortium analysis (Han et al., 2008; Sulem et al., 2007), we had seen association to a nearby SNP, but with the strength of the association to rs12203592 and later functional work (Praetorius et al., 2013, showed that this was the key variant), we observed a most interesting flip-flop of the association with mole count (Duffy et al., 2010), depending on whether we looked at raised moles or a flat moles in the BTNS twins, which also led to changes in the allele associated with high total mole count depending on age when we compared the young twins to either their own parents, or to other populations of adults. Figure 1 shows the GWAS for total nevus count in 12-year-old twins from BTNS as of January 2020 (from analysis by Gu Zhu). The two most significant associations are over the *IRF4* (chr 6) and *MTAP* (chr 9) genes, but *PLA2G6* on chr 22 still does not get a look-in — further evidence of age heterogeneity.

https://doi.org/10.1017/thg.2020.38 Published online by Cambridge University Press

One culmination of this mole work is the big consortium paper we led that included 3261 children and 2248 of the parents from the BTNS among a total of 52,000 individuals from around the world (Duffy et al., 2010, 2018). By combining these data with the results of an earlier melanoma meta-analysis led by our group (Law et al., 2015), we were able to implicate 30 genes as controlling mole count and melanoma risk, most affecting both traits equally, but some just for mole count (e.g., *KITLG*, a gene already known as a pigmentation locus), and a few just for melanoma. Given the fact that we actually counted moles (many of the other studies relied on questionnaire self-report), we were confident our contribution to the study power was much higher than raw numbers might suggest. I have just highlighted a few of the many mole and melanoma associated BTNS papers, but these make up only a small fraction of the total number of papers arising from this study.

The studies of eye color described above foreshadowed Nick's involvement in studies of eye traits more generally. In the mid 2000s, in collaboration with David Mackey, the BTNS families were phenotyped for a wide range of quantitative traits of relevance to eye health. While the first of these studies employed the linkage approach, the strategy really began to yield genes when the GWAS was applied. The first successful GWAS internationally was on the eye disease, age-related macular degeneration (AMD) in 2005. Studies involving Nick's twins for a range of eye traits followed soon after. As foreshadowed by AMD, the genetic architecture of eye traits proved to be more tractable than most other complex traits. For example, the first GWAS on the eye trait optic disc area found common alleles that explained up to 3% of the variance (Macgregor et al., 2010), more than 10 times the effect size seen for traits such as body mass index or height. Twins also formed the basis for many subsequent studies of eye disease, including studies on myopia (Hysi et al., 2010; Law et al., 2015), keratoconus and related traits (Lu et al., 2013), and on glaucoma risk factors including intraocular pressure (Hysi et al., 2014). By 2018, expanded sample sizes meant that hundreds of genes had been uncovered for myopia (Tedja et al., 2018) and intraocular pressure (MacGregor et al., 2018).

Although the BTNS sampling frame was children, there are now excellent examples where the endophenotype (disease risk factor) approach has borne fruit, with important consequences for diseases in later life. As noted above, in the case of mole count, ever larger GWASs, frequently comprising large numbers of individuals too young to be personally affected by cancer, have yielded many genes that were subsequently shown to influence melanoma risk. In the case of eye disease, the same genes that influence a person's risk of intraocular pressure in early to mid life turned out to be excellent predictors of glaucoma risk in later life (MacGregor et al., 2018). Recent work in this space has illustrated the potential for gene mapping findings to be translated to disease prevention, for example, in glaucoma, it was recently shown that by combining endophenptype data from healthy cohorts such as BTNS with data on glaucoma case–control cohorts, that it was possible to derive glaucoma-specific genetic risk scores (Craig et al., 2020). These genetic risk scores are showing promise in determining who is likely to be at highest risk of early-onset glaucoma, an exciting outcome given glaucoma is eminently preventable if detected and treated early.

Nick's boundless enthusiasm for setting up genetic studies has enabled advances across a wide range of diseases and traits. BTNS is an exemplar of Nick's ability to set up and capitalize on twin data. In this article, we have only covered nevus count, eye disease and related traits, although BTNS has enabled research on a much wider range of traits. As well as traits from questionnaires and nurse measurements, in collaboration with Peter Visscher and others, BTNS was characterized for gene expression and methylation (Powell et al., 2012), further expanding the scope of a study originally funded to examine nevus count. Nick's twin data have formed the foundation of a vast number of publications. Indeed, the sheer number of resultant articles from a single scientist has prompted incredulity from some commentators (Ioannidis et al., 2018). In Nick's case, his publication count far in excess of 1000 does not exaggerate his impact — rather, it is a reflection of his outstanding work over an extended period, building and leveraging twin cohorts to advance scientific knowledge across a broad range of scientific endeavors.

References

Box, N. F., Duffy, D. L., Chen, W., Stark, M., Martin, N. G., Sturm, R. A., & Hayward, N. K. (2001). MC1R genotype modifies risk of melanoma in families segregating CDKN2A mutations. *American Journal of Human Genetics*, *69*, 765–773.

Box, N. F., Wyeth, J. R., O'Gorman, L. E., Martin, N. G., & Sturm, R. A. (1997). Characterization of melanocyte stimulating hormone receptor variant alleles in twins with red hair. *Human Molecular Genetics*, *6*, 1891–1897.

Craig, J. E., Han, X., Qassim, A., Hassall, M., Cooke Bailey, J. N., Kinzy, T. G., & MacGregor, S. (2020). Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. *Nature Genetics*, *52*, 160–166.

Duffy, D. L., Iles, M. M., Glass, D., Zhu, G., Barrett, J. H., Höiom, V., & Montgomery, G. W. (2010). IRF4 variants have age-specific effects on nevus count and predispose to melanoma. *American Journal of Human Genetics*, *87*, 6–16.

Duffy, D. L., Montgomery, G. W., Chen, W., Zhao, Z. Z., Le, L., James, M. R., & Sturm, R. A. (2007). A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. *American Journal of Human Genetics*, *80*, 241–252.

Duffy, D. L., Zhu, G., Li, X., Sanna, M., Iles, M. M., Jacobs, L. C., & Martin, N. G. (2018). Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. *Nature Communications*, *9*, 4774.

Falchi, M., Bataille, V., Hayward, N. K., Duffy, D. L., Bishop, J. A. N., Pastinen, T., & Spector, T. D. (2009). Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. *Nature Genetics*, *41*, 915–919.

Han, J., Kraft, P., Nan, H., Guo, Q., Chen, C., Qureshi, A., & Hunter, D. J. (2008). A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genetics*, *4*, e1000074.

Hysi, P. G., Cheng, C.-Y., Springelkamp, H., Macgregor, S., Bailey, J. N. C., Wojciechowski, R., & Aung, T. (2014). Genome-wide analysis of multi-ancestry cohorts identifies new loci influencing intraocular pressure and susceptibility to glaucoma. *Nature Genetics*, *46*, 1126–1130.

Hysi, P. G., Young, T. L., Mackey, D. A., Andrew, T., Fernández-Medarde, A., Solouki, A. M., & Hammond, C. J. (2010). A genome-wide association study for myopia and refractive error identifies a susceptibility locus at 15q25. *Nature Genetics*, *42*, 902–905.

Ioannidis, J. P. A., Klavans, R., & Boyack, K. W. (2018). Thousands of scientists publish a paper every five days. *Nature*, *561*, 167–169.

Law, M. H., Bishop, D. T., Lee, J. E., Brossard, M., Martin, N. G., Moses, E. K., & Iles, M. M. (2015). Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. *Nature Genetics*, *47*, 987–995.

Lu, Y., Vitart, V., Burdon, K. P., Khor, C. C., Bykhovskaya, Y., Mirshahi, A., & Wong, T. Y. (2013). Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus. *Nature Genetics*, *45*, 155–163.

Macgregor, S., Hewitt, A. W., Hysi, P. G., Ruddle, J. B., Medland, S. E., Henders, A. K., & Mackey, D. A. (2010). Genome-wide association identifies ATOH7 as a major gene determining human optic disc size. *Human Molecular Genetics*, *19*, 2716–2724.

https://doi.org/10.1017/thg.2020.38 Published online by Cambridge University Press

MacGregor, S., Ong, J.-S., An, J., Han, X., Zhou, T., Siggs, O. M., & Hewitt, A. W. (2018). Genome-wide association study of intraocular pressure uncovers new pathways to glaucoma. *Nature Genetics*, *50*, 1067–1071.

McGregor, B., Pfitzner, J., Zhu, G., Grace, M., Eldridge, A., Pearson, J., & Martin, N. G. (1999). Genetic and environmental contributions to size, color, shape, and other characteristics of melanocytic naevi in a sample of adolescent twins. *Genetic Epidemiology*, *16*, I40–I53.

Palmer, J. S., Duffy, D. L., Box, N. F., Aitken, J. F., O'Gorman, L. E., Green, A. C., & Sturm, R. A. (2000). Melanocortin-1 receptor polymorphisms and risk of melanoma: is the association explained solely by pigmentation phenotype? *American Journal of Human Genetics*, *66*, 176–186.

Powell, J. E., Henders, A. K., McRae, A. F., Caracella, A., Smith, S., Wright, M. J., & Montgomery, G. W. (2012). The Brisbane Systems Genetics Study: Genetical genomics meets complex trait genetics. *PloS One*, *7*, e35430.

Praetorius, C., Grill, C., Stacey, S. N., Metcalf, A. M., Gorkin, D. U., Robinson, K. C., & Steingrimsson, E. (2013). A polymorphism in IRF4 affects human pigmentation through a tyrosinase-dependent MITF/TFAP2A pathway. *Cell*, *155*, 1022–1033.

Sturm, R. A., Duffy, D. L., Box, N. F., Chen, W., Smit, D. J., Brown, D. L., & Martin, N. G. (2003). The role of melanocortin-1 receptor polymorphism in skin cancer risk phenotypes. *Pigment Cell Research*, *16*, 266–272.

Sturm, R. A., Duffy, D. L., Zhao, Z. Z., Leite, F. P. N., Stark, M. S., Hayward, N. K., Martin, N. G., & Montgomery, G. W. (2008). A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. *American Journal of Human Genetics*, *82*, 424–431.

Sulem, P., Gudbjartsson, D. F., Stacey, S. N., Helgason, A., Rafnar, T., Magnusson, K. P., & Stefansson, K. (2007). Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature Genetics*, *39*, 1443–1452.

Tedja, M. S., Wojciechowski, R., Hysi, P. G., Eriksson, N., Furlotte, N. A., Verhoeven, V. J. M., & Klaver, C. C. W. (2018). Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. *Nature Genetics*, *50*, 834–848.

Valverde, P., Healy, E., Jackson, I., Rees, J. L., & Thody, A. J. (1995). Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nature Genetics*, *11*, 328–330.

Valverde, P., Healy, E., Sikkink, S., Haldane, F., Thody, A. J., Carothers, A., … Rees, J. L. (1996). The Asp84Glu variant of the melanocortin 1 receptor (MC1R) is associated with melanoma. *Human Molecular Genetics*, *5*, 1663–1666.

Zhu, G., Duffy, D. L., Eldridge, A., Grace, M., Mayne, C., O'Gorman, L., & Martin, N. G. (1999). A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: A maximum-likelihood combined linkage and association analysis in twins and their sibs. *American Journal of Human Genetics*, *65*, 483–492.

Zhu, G., Evans, D. M., Duffy, D. L., Montgomery, G. W., Medland, S. E., Gillespie, N. A., & Martin, N. G. (2004). A genome scan for eye color in 502 twin families: Most variation is due to a QTL on chromosome 15q. *Twin Research*, *7*, 197–210.

https://doi.org/10.1017/thg.2020.38 Published online by Cambridge University Press