Nicholas Gordon Martin

Table of Contents

PREFACE

Dear Nick,

In compiling and editing this compendium of commentaries and anecdotes about how you have influenced (and continue to influence) the field of complex trait genetics, a number of consistent themes and highlights emerged:

> Your founding of the discipline of human complex trait genetics in Brisbane and Australia more broadly
>
> Your founding of the Australian Twin Registry and the Brisbane twin resource
>
> Your generosity of spirit with respect to scientific ideas, authorship and your unconditional sharing of vast data resources amongst collaborators
>
> Your intellectual and financial generosity as a supervisor and mentor
>
> Your incisive mind and intellectual curiosity
>
> Your lasting contributions to the development of statistical methods to model human resemblance
>
> Your leadership and major contributions in the world of GWAS
>
> Your tongue in cheek (at least we think) political incorrectness

In the following pages, we and others have tried to document the myriad ways in which you have influenced the field of complex trait genetics and changed the lives of the many students, employees, mentees, colleagues and collaborators you have worked with.

Given the hundreds of scientists you have collaborated with over your career, inevitably there will be some whom we have missed, who will have wanted to contribute their own anecdotes and stories about how you have influenced their careers. Indeed during the course of compiling this tome, many more individuals came forward wishing to contribute, and we have done our best to include them here. For those whom we have omitted, we apologise and take full responsibility for the oversight. Indeed, the fact that this compendium is so long and contains so many contributions is testament to influence you have had!

Happy 71st year Nick. We do hope you enjoy reading the multitude of ways in which you have influenced the field of complex trait genetics and made major positive impacts on the lives and scientific careers of so many. Indeed your legacy is not only a broad list of major scientific achievements in the field and the wonderful resources you have founded, but your intellectual F1s and F2s (some of them genetically related!), many of whom have returned to Brisbane after extended stays abroad, and are now leaders in the field in their own right.

Compiling and editing this tome has certainly been "a good exercise for us".

David Evans, Sarah Medland and Nathan Gillespie, on behalf of the many contributors.

February 2020, Brisbane Australia

List of Contributors

Meike Bartels *Free University Amsterdam*

Tim Bates *University of Edinburgh*

Dorret Boomsma *Free University Amsterdam*

Cynthia Bulik *University of North Carolina*

Enda Byrne *University of Queensland*

Georgia Chenevix-Trench *QIMR Berghofer Medical Research Institute*

Stacey Cherny *Tel Aviv University*

Lucía Colodro-Conde *QIMR Berghofer Medical Research Institute*

Baptiste Couvy-Duchesne *QIMR Berghofer Medical Research Institute*

Eske Derks *QIMR Berghofer Medical Research Institute*

David Duffy *QIMR Berghofer Medical Research Institute*

Lindon Eaves *Virginia Commonwealth University*

David Evans *University of Queensland*

Nathan Gillespie *Virginia Commonwealth University*

Pete Hatemi *Penn State University*

Anjali Henders *University of Queensland*

John Hewitt *University of Colorado Boulder*

Ian Hickie *University of Sydney*

Daniel Hwang *University of Queensland*

Jaakko Kaprio *University of Helsinki*

Matthew Keller *University of Colorado Boulder*

Martin Kennedy *University of Otago*

Penelope Lind *QIMR Berghofer Medical Research Institute*

Michelle Luciano *University of Edinburgh*

Stuart MacGregor *QIMR Berghofer Medical Research Institute*

Hermine Maes *Virginia Commonwealth University*

Hilary Martin *Sanger Institute*

Sarah Medland *QIMR Berghofer Medical Research Institute*

Christel Middeldorp *University of Queensland*

Grant Montgomery *University of Queensland*

Ben Neale *The Broad Institute*

Mike Neale *University of Colorado Boulder*

Dale Nyholt *Queensland University of Technology*

Danielle Posthuma *Free University Amsterdam*

Shaun Purcell *Harvard Medical School*

Pak Sham *University of Hong Kong*

Wendy Slutske *University of Missouri*

Rick Sturm *University of Queensland*

Brad Verhulst *Texas A&M University*

Karin Verweij *Amsterdam University Medical Centers*

Peter Visscher *University of Queensland*

Tracey Wade *Flinders University*

John Whitfield *QIMR Berghofer Medical Research Institute*

Naomi Wray *University of Queensland*

Jian Yang *University of Queensland*

Gu Zhu *QIMR Berghofer Medical Research Institute*

Brendan Zietsch *University of Queensland*

Martin NG & Martin PG (1975). The inheritance of scholastic abilities in a sample of twins. I Ascertainment of the sample and diagnosis of zygosity. Ann Hum Genet Lond, 39, 213-218. (cites = 195)

Martin NG (1975). The inheritance of scholastic abilities in a sample of twins. II Genetical analysis of examination results. *Ann Hum Genet*, 39, 219-229. (cites = 37)

(Nick and Georgia's wedding 1983)

Birmingham and Beyond.

Lindon Eaves

The destinies of mentors and students are closely intertwined. I am blessed by and envious of the accomplishments of both over the last half-century. It is a great privilege to pay tribute to Nick.   He is hard to keep up with. He walks faster than me. He has far greater insight about the worldly twists of academic lives and influence, he is more knowledgeable, more energetic, more up-to-date, more generous and more passionate about science than most people I have known.   I remember walking quickly with him through the streets of Rome during a Twin Congress, commenting on the significance of the ancient designation "S.P.Q.R" on the metal drain covers in the side-walks. "Senatus Populusque Romano", I said. "No", said Nick, "Senat<u>U</u> Populoque Romano. Ablative not nominative."

Nick taught me much more than science.   A few hours between talks at my first Twin Congress (Washington, D.C., 1977) found me following him breathlessly around the museums and galleries on the Mall, culminating in a visit to the Air and Space Museum that left me in awe of such a cathedral to courage and ingenuity. As a poor British post-doc, I could never have afforded the trip to America had Nick not dictated the letter he insisted I send to the organizers begging for a paid invitation. This led to a sabbatical a year later and, ultimately, to emigration from Oxford to Richmond with my wife, two children and the family cat.   Nick's wedding to Georgia in Richmond a few years later was the occasion of my becoming licensed to perform weddings in the Commonwealth of Virginia so that I could assist.

In the early 1970's, I was a keen new post-doc with the MRC research program in Psychogenetics at Birmingham. The program was directed jointly by John Jinks, F.R.S., Chair of Genetics and Peter Broadhurst, Chair of Psychology.   In collaboration with Sir Kenneth Mather, F.R.S., the previous chair, Jinks had established Birmingham as an international center for "Biometrical Genetics" in the study of polygenic effects on continuous variation across a range of model systems in flowering plants, micoorganisms, fruitflies, rodents and, most recently, humans. The program in psychogenetics had its roots in earlier previous collaboration between Jinks and Broadhurst on behavioral development in the laboratory rat.   In the late 60's, David Fulker pioneered early classical papers on the biometrical genetic analysis of behavior in *Drosophila*, rats and humans.

I had read some of this early work in 1966 as an undergraduate in Genetics. Two years later, John Jinks generously agreed to becoming my pre-doctoral mentor in his department after my detour through seminary ended in Fall, 1968.   Early in 1968 I had met with Jinks and Broadhurst to discuss possibilities of graduate research in human behavioral genetics.   They presented me with a preprint of the Jinks and Fulker landmark "A Comparison of the Biometrical Genetical, MAVA and Classical Approaches to the Analysis of Human Behavior" (Pychol. Bull, **73,** 311-349, 1970). This paper still represents one of the intellectual landmarks in our field for its introduction of human behavioral geneticists to the power of a model-fitting approach to guide the process of choosing among multiple conceptions of the underlying causes of human variation and assigning numerical estimates to unknown parameter values.

I was leafing through my daily pile of line-printer output one morning in the early 70's when the phone rang. "G'day", I heard. "Nick Martin here." "Uh?" I thought. "Did you read that paper I sent you?" he asked, referring to a preprint of a paper on the inheritance of scholastic abilities in a sample of Australian Twins coauthored with his father, P.G. Martin

(Annals of Hum. Genet.,**39**,231-218, 1975)   "D[ur]n!" I thought, not having read it. "Ahem!" I coughed, trying to hide my embarrassment. "Where are you?" I asked, changing the subject. "Oxford", he replied, "Can I come and see you?"   So Nick came to Birmingham, becoming my first graduate student, life-time friend, colleague, eminent scholar and inspiration.

The next decade was a time of energetic transformation and clarification.   Most of the basic ideas were already there in the publications of distinguished colleagues around the world but the pieces needed gathering together and sharpening.   Three problems emerged as most pressing in the next decade.

1. Extension of the "model-fitting" approach of Jinks and Fulker to the multivariate case.   The power of biometrical genetics tended to focus on specific single model variables such as final height and time of first flowering in Nicotiana rustica, number of sternopleural chaetae in Drosophila, growth rate in Aspergillus sp. The study of human behavior, by contrast, was often inherently multivariate building upon the early psychometric studies of the structure of multiple abilities and personality. Several pioneers (e.g. Loehlin and Vandenberg, 1968) attempted to answer the question of how genes and environment imparted structure to the pattern of covariation between multiple variables, especially human abilities.   Most of these early attempts were modifications of existing methods for multivariate analysis, such as factor analysis of estimated components of variance and covariance or attempts to squeeze twin data into the multivariate analysis of variance.   Was it possible to extend the heuristic power of Jinks and Fulker's model-fitting approach to the multivariate case?   On a visit to our collaborator Hans Eysenck at the Institute of Psychiatry in London, the problem was outlined to the late Owen White, co-developer of the Promax algorithm for oblique factor rotation. "Sounds like you need to look at Joreskog (*Psychometrika,* **43,** *443-447,* 1978) on the analysis of covariance structures.   Maybe look at 'LISREL'".   Owen was the unheralded inspiration for the next step. Unfortunately, the initial versions of LISREL, could not quite handle the multiple group problem inherent in kinship analyses but we were able to figure out how to write our own crude FORTRAN IV programme and apply it to a small set of multivariate twin data on primary mental abilities generously supplied by John Loehlin and Steven Vandenberg (Martin and Eaves, *Heredity,* **38,** 79-95).   Plant and fruitfly colleagues, used to the fine dissection of genetic effects in complex breeding experiments, were skeptical of these crude attempts. One "drosophilist" colleague remarked over lunch one day: "Hm! I have enough problems doing the genetics of one variable let alone spending time trying to analyze five."

The program, laboriously coded on punched cards, took all night to run on the Birmingham University 240K, KDF-9 main-frame computer but seemed to work and give sensible answers. These days, a similar analysis in Mx probably runs in one second on a $500 lap-top.   In the 40 years since, the computer revolution has made it possible for other investigators to extend, teach and apply this basic approach to human quantitative data. Others will write of the twin "workshops", taught initially in Leuven using updated more flexible versions of LISREL in the presence of Karl Joreskog, and of Mike Neale's persistent work on the development of Mx to allow the unprecedented flexibility of models implying non-linear parameter constraints.

Like me, Nick experienced the Birmingham course in biometrical genetics.   It combined magisterial lectures with exhausting day-long, hands-on analytical sessions where we had to compute generation means and variances as a prelude to weighted least-squares estimation of biometrical model parameters. The nadir was the requirement of inverting by hand the 4x4 information matrix of additive, dominant and environmental components of

means and variances on an electro-mechanical calculator. But everyone who had to do it learned to look at and think about the raw numbers. Every single plant or fruit-fly mattered. For Nick, as for many who sat through those long days, the classes were a never-forgotten model of teaching.

In the conference lobby at another Twin meeting, after a depressing series of papers with little common conceptual or analytical thread, Nick energetically urged the possibility of using the Birmingham approach to teach new generations of researchers. Thus, were born the first Leuven NATO workshops, under the hospitable eye of Bob Vlietinck and Robert Derom. Later the Leuven workshops evolved, with NIMH support, into the current series of Boulder workshops hosted by John DeFries, John Hewitt and their colleagues at the I.B.G. On many occasions, informal gatherings of faculty at the back of the room were the treasured occasion for exploring and discussing new problems.

2. Towards better models for genes and environment in human behavior. The early models for genetic effects were embarrassingly simple and due largely to the genius of Ronald Fisher. Our fungal, fly and fruit-fly colleagues knew that the effects of genes were far more subtle than Fisher's basic additive and dominant components of variance. Our colleagues in psychology were skeptical of the simple partition of the environment into effects shared and not shared by family members. The elements of more subtle models were already recognized in plant and animal studies with the recognition that different genes may be expressed in different environments and that part of the environmental variation between individuals is a function of the genotypes of their relatives, spouses and peers (the "genetic environment").

The issue of how best to integrate a biometrical-genetic approach to genetic effects, modeled basically on the pioneering work of Fisher (*Trans. Roy. Soc. Edin.,* **52,** 399-433, 1918) and Mather (*Biometrical Genetics,* London, Methuen,1949) with a mathematical formulation that allowed for the non-genetic interaction among family members was a source of much controversy, even acrimony at times. Different groups applied different numerical approaches to data. They disagreed about the relevance of different theoretical assumptions about genes, environment and mate selection and even differing traditions of notation for genetic variation. At times, the academic dialogue was even described in terms of "schools" holding fast to different views of what is worth doing. Controversy about genes and environment in that climate was further compounded by disagreement about the social and political implications of behavioral genetics and the role of single gene models derived from medical genetics to the complexities of quantitative human traits. The biometrical genetic heuristic was articulated clearly in discussion of a conference paper in Eaves (*J. Roy. Statist. Soc. A,* **140**, 324-355, 1977). When a senior skeptic questioned why John Jinks' group weren't trying to look for the individual genes of large effect, Jinks responded testily (and presciently) on the basis of his life's work so far: "The number of genes is directly proportional to the industry of the investigator".

The principal conflict of the 70's arose between those who had learned the ways of Birmingham which followed the intellectual descendants of Fisher in emphasizing the richness of genetic variation and gene action and those, following Newton Morton and his colleagues (e.g. Rao, DC, Morton, NE, Yee S. *Am. J. Hum. Genet,* **26,** 331-359, 1976), who had rediscovered the potential of path analysis to recover some of the major sources of non-genetic inheritance and spousal resemblance. It was a minefield of strongly held loyalties that led to some caustic exchanges. I still remember Jinks' withering skepticism when he came into the lab and saw me drawing a path model on a piece of computer paper. Also, I recall a passing conversation

with Mather in the corridor when he noted his surprise that a paper I had published referred to the additive genetic variation as Douglas Falconer's "V_A" rather than his "$^1/_2D_R$".   These were all first fumblings of people who were trying to do their best and "get it right" in doing justice to the special problems associated with family resemblance in human behavior. It is almost embarrassing to look back at a summary published by a group of us, including Nick (Eaves et al, *Heredity,* 1978).   There is almost no mention of assortative mating because biometrical genetics had little to say about it.   Indeed, most of what had been said by others used path analysis made untested assumptions about the underlying process of mate selection. How we wished for more I.Q. points! In the same year, John Rice and his colleagues (Rice et al, *Am. J. Hum. Genet.* **30,** 618-643, 1978) published a landmark attempt to integrate Fisher's (1918) model for polygenic inheritance, assortative mating with the insight of Cavalli-Sforza and Feldman (*Am. J. Hum. Genet.* **25,** 618-637, 1973) that cultural transmission implied direct transmission from parental to child phenotypes.   There were still a few more steps. It was a privilege to spend two years in Oxford as Andrew Heath's doctoral advisor.   One of our burning questions was how to extend the model of Rice et al to include both phenotypic assortment and social homogamy.   I certainly didn't have a clue. I remember making a few comments to Andrew about my frustration during a tea-break (see "Importance of Tea Break" below) in the Psychology common-area. Next morning, I arrived at the office to find Andrew had figured it all out beautifully and attempted patiently to explain it to my flagging comprehension. This would not be the last time.

3. Design, Sample Size and Data.

By the end of the 70's it seemed that many important statistical and numerical questions about how to develop and test quite complex models for human behavior had been resolved. These developments allowed a lot to the enormous growth in computer power and development of efficient software for numerical optimization embodied for example in that developed by the Numerical Algorithms Group from which we readily borrowed.   What we lacked were the data. What kinds of data? How much data? Who was going to collect it? Who was going to pay for it?   As long as the focus of behavior genetics lay in estimating heritability or testing for the non-genetic correlation between relatives, family studies before the 1970's were typically small, perhaps 10's or a few 100 relative pairs.   Sampling errors, when they were computed were large and the questions very simple.   The influence of our colleagues in biometrical genetics including Mike Kearsey and Brian Barnes introduced us to the value of computer simulation of sample size and experimental design to resolve different components of the biometrical genetical model.   Human application of this was exemplified in the paper by Martin et al (*Heredity,* **40,** 97-116, 1978) on the power of the classical twin study. These early simulations heightened our awareness that samples in the 1,000's or larger were a prerequisite for reliable inference.

During the early 80's, after Nick and Andrew moved to Virginia, this basic approach was extended to the more complex designs that had been generated by the extension of the classical twin paradigm to include the children of MZ and DZ twins (Nance and Corey, *Genetics, 83,* 811-825, 1976) to explore the effects of the maternal genotype on child development (see also Haley et al., *Heredity,* **46,** 227-238, 1981).   We were also absorbed by the information that the spouses of twins might yield about the social and phenotypic effects of mate selection (Heath and Eaves, *Behav. Genet.***15,** 15-30, 1985).   The arrival of Ken Kendler at the Department of Psychiatry was also an important stimulus to implementing a dream about resolving the possible correlation and interaction of genes and environment by incorporating intensive individual environmental measures in a large psychiatric twin study.   Among other studies spawned in this period were Dr. Kendler's long-standing series of adult twin studies,

the Virginia 30,000 study of the extended kinships of twins and the Virginia Twin Study of Adolescent Behavior Development.   Elements of these studies have been transformed into further major studies pioneered by Nick and Andrew after their departure to prestigious appointments in Brisbane and St. Louis respectively.

4. The importance of tea-break.

Effective science thrives in the crucible of collegiality.   All of us who worked in Birmingham at that time remember the twice-daily rumble of the tea trolley pushed from the departmental kitchen to the genetics library where the faculty and students gathered, surrounded by racks of periodicals.   In those days, genetics crossed many disciplines from statistics to cytogenetics and biochemical genetics. Molecular genetics was nascent.   It is impossible to do justice to those tea-breaks. The range of characters, social and political values, scientific depth was astonishing and moving.   Exchanges were sometimes caustic. Many were the times when questions would arise that led to follow-up side bars where we got to pick the brains of colleagues who knew more than we did. After our arrival in Richmond, space was limited, so we occupied an unused wet lab which soon became littered with piles of paper. Largely under Nick's influence we continued the tradition of tea-break, often gathered round the black board, sometimes adjourning for a sandwich lunch at the "Skull and Bones", a darkly named restaurant associated with the University Hospital. There ballpoint and napkin took the place of blackboard and chalk.

The expansion of molecular genetics eventually led to our eviction from the wet lab and the removal of our piles of line-printer output, reprints etc. to a more appropriate setting for family-based research.   But it turned out we would not be forgotten quite so easily.   When the space was remodeled it was discovered that the drain from the sink had been blocked by a lasting accumulation of lapsang souchong leaves. Nick's special tea.

Lindon Eaves,
Richmond, VA,
February 2020

(Lindon Eaves and Nick Martin, Egmond Aan Zee 2004)

Chenevix-Trench G, Jones K, Green AC, Duffy DL, Martin NG (1992). Cleft lip with or without cleft palate: associations with transforming growth factor alpha and retinoic acid receptor loci. *Am J Hum Genet*, 51(6), 1377-85. (cites = 242)
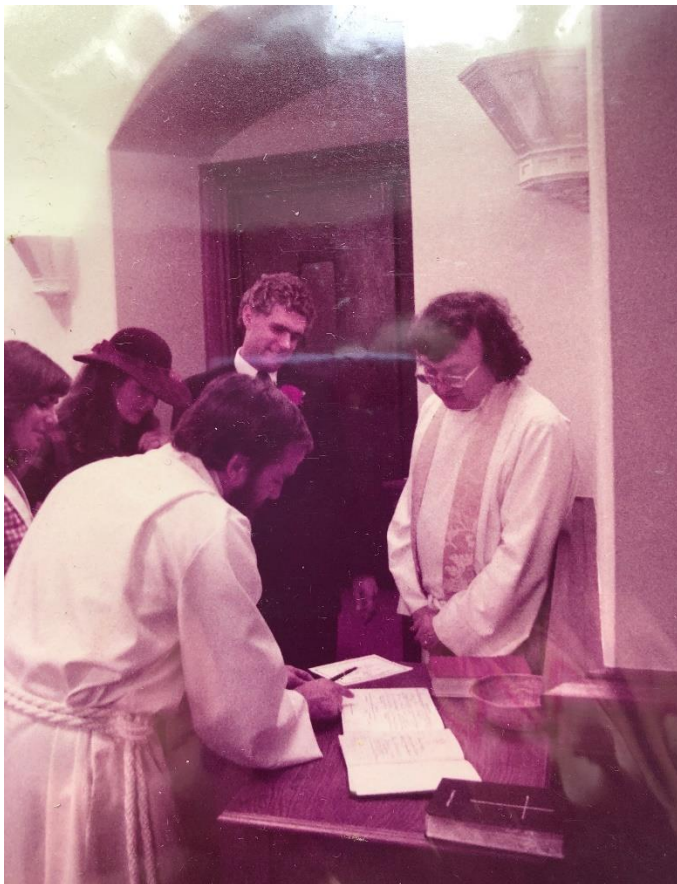
(Nick and Georgia's wedding 1983)

Nicholas Gordon Martin

Georgia Chenevix-Trench

I've been asked to write about 'the significance of a particular aspect of Nick's work on the field of genetics'. I know that others will talk about his contributions to understanding the genetic basis of a wide range of traits from DZ twinning to drug taking, drinking, depression, and DNA methylation, but I am perhaps best qualified to comment on his influence on one naïve PhD student of genetics who turned into his Head of Department. I should point out that probably the only reason he himself is not the HoD is that he refuses to attend any meetings longer than an hour, thereby very efficiently ruling himself out of any administrative role which he would see as merely a distraction from doing what he loves – setting up collaborations, nursing along projects, teaching students and red-penning manuscripts.

I was a second year PhD student at the Medical College of Virginia when the Reverend Dr Lindon Eaves told us that a mad, bad and dangerous Australian was coming to visit and we should all be very careful. And so, dear Reader, Lindon married us, two years later. In retrospect, I had no idea what I was doing, starting down an academic pathway. I am the only scientist in my family, had no idea about the life of a research scientist, and had a PhD supervisor who had a 'light touch', to say the least. Nick was the first person I'd met who lived and breathed genetics, who never questioned that research was all that he wanted to do (indeed all that anyone would want to do), and who was confident enough to know that if that is what you want, then it will surely happen. For these reasons and more, he always assumed that we'd be doing this together, and that we'd share the responsibilities of house and family. I am sure it is profoundly irritating to those who might be tempted to dismiss him as a MCP (that is, a Male Chauvinist Pig - does that term even exist anymore?) when he injects some inflammatory comment to get a rise, to know that he fights to get *into* the kitchen at night, is better acquainted with the vacuum cleaner than I am, started solo parenting when Hilary was six months old and I went off to a conference, and won the nappy-changing and hanging-up-washing (with baby on hip) contests at her 1st birthday party.

I come from a long line of writers, so fancied myself as pretty competent at sentence construction at least. But I am sure I was not the first person he reduced to tears by covering the first manuscript I wrote with red-penned comments, although I did realise that he was (usually) right and listened through gritted teeth until he declared that I had 'graduated' from his writing school. Before I finished my PhD we had decided that we wanted to move to Australia, even though when we first met I told him that I had no interest in going to Australia, and that if someone gave me a ticket I'd trade it in. In 1985 there were so few jobs, anywhere in the world, in statistical genetics that beggars couldn't be choosers and when one came up in Brisbane, he had to take it once they had thrown in a post-doc position for me. Back then, Brisbane was a total backwater scientifically (and indeed in other ways, being not much more than a country town) so when we told people we were moving there, I think they thought we must have a criminal record, and might as well move to Vladivostok. Such is Nick's incredible optimism, and love for Australia, that he took on the challenge and we came on a one-way ticket to a town I'd never been where the houses are on stilts, to a post-doc position with no specified supervisor, to a country 'of droughts and flooding rains'. The timing was right, because a very right-wing State government had left the coffers full, and a long-term Labour government was very supportive of research. But mostly it was Nick's energy and force of personality that turned Brisbane into a world class centre of complex trait genetics. The list of

people who either trained with Nick, or became regular visitors, or are his academic 'grandchildren' is a veritable Who's Who of the field now, and I am sure into the future.

I helped him get his lab running when we first arrived, and we worked together on the genetics of cleft lip and palate. Our candidate SNP study on 117 cases and 113 controls, published in the *American Journal of Human Genetics* in 1992, was quite a landmark association study for a complex trait and has 242 citations! He isn't entirely one track minded in his professional life though: when Hilary admitted to wanting to be a geneticist he said she could be a statistical geneticist, a quantitative geneticist or a genetic epidemiologist. And he has even come around to Felicity's career choice – she calls herself the 'black sheep' of the family as a medical doctor because he knows she is so good at it.

Nick has an incredible knack of seeing the best in those who might otherwise be thought of as (and indeed, often proved to be) ill qualified for the job (like the taxi driver he once gave a job to because he said that he thought that melanoma 'was all about the melacortin 1 receptor'), and the Iraqi refugee (an old friend from his PhD days, then a professor in Mosul) who had never used a keyboard, much less the internet or done a PCR reaction. There is nothing he likes as much as filling the house with students and post-docs, putting on some Schubert or Ali Farka Toure, pouring a stiff gin and tonic or two, cooking a roast lamb with potatoes done to perfection in duck fat (ssshhh – don't tell the vegetarians) and winding people up through a fierce (and only semi-serious) political argument. One of the best such parties ended with the irreplaceable Leena Peltonen saying 'Neek, I lurve you, every time I meet you I lurve you, but you are stoopid, stoopid and stoopid'!

Despite his devotion to genetics, Nick epitomizes 'work hard and play hard', a culture that is encouraged in Australia. Anyone who has ever doubted his role on the more than 1300 papers he is an author on should see the red pen marks all over our bed sheets – most of them he reads from 5-6am. He has been to work almost every Sunday of his working life but in return we take off all of January and have just returned from a six week holiday in Africa.

It is always difficult to imagine 'what if' but I am confident that if we hadn't seen that advertisement in *Nature* for a genetic epidemiologist in Brisbane, one snowy Sunday in Richmond, the field would be a very different place and none of you would have ever heard of the Queensland Institute of Medical Research.

(Nick as a boy)

Martin NG, Olsen ME, Theile H, El Beaini JL, Handelsman D, Bhatnagar AS (1984). Pituitary-ovarian function in mothers who have had two sets of dizygotic twins. *Fertility and Sterility*, 41(6), 878-880. (cites = 53)

Martin NG, Robertson DM, Chenevix-Trench G, de Kretser DM, Osborne J, Burger HG (1991). Elevation of follicular phase inhibin and luteinizing hormone levels in mothers of dizygotic twins suggests nonovarian control of human multiple ovulation. *Fertility and Sterility* 56(3), 469-474. (cites = 39)

Lambalk CB, Boomsma DI, Boer L de, Koning CH de, Schoute E, Popp-Snyders C, Schoemaker J (1998). Increased levels and pulsatility of Follicle-Stimulating Hormone in mothers of hereditary dizygotic twins. *J Clinical Endocrinology and Metabolism*, 83, 481-486. (cites = 60)

Martin NG, Shanley S, Butt K, Osborne J, O'Brien G (1991). Excessive follicular recruitment and growth in mothers of spontaneous dizygotic twins. *Acta Geneticae Medicae et Gemellologiae,* 40(3-4), 291-301. (cites = 34)

Lewis CM, Healey SC, Martin NG (1996). Genetic contribution to DZ twinning. *Am J Medical Genetics*, 61(3), 237-246. (cites = 44)

Meulemans WJ, Lewis CM, Boomsma DI, Derom CA, Berghe H van den, Orlebeke JF, Vlietinck RF, Derom RM (1996). Genetic modeling of dizygotic twinning in pedigrees of spontaneous dizygotic twins. *Am J Medical Genetics*, 61, 258-263. (cites = 50)

Painter JN, Willemsen G, Nyholt D, Hoekstra C, Duffy DL, Henders AK, Wallace L, Healey S, Cannon-Albright LA, Skolnick M, Martin NG, Boomsma DI, Montgomery GW (2010). A genome wide linkage scan for dizygotic twinning in 525 families of mothers of dizygotic twins. *Hum Reprod*, 25(6), 1569-80. (cites = 34)

Gajbhiye R, Fung JN, Montgomery GW (2018). Complex genetics of female fertility. *NPJ Genom Med*, 3, 29. (cites = 4)

Mbarek H, Steinberg S, Nyholt DR, Gordon SD, Miller MB, McRae AF, Hottenga JJ, Day FR, Willemsen G, de Geus EJ, Davies GE, Martin HC, Penninx BW, Jansen R, McAloney K, Vink JM, Kaprio J, Plomin R, Spector TD, Magnusson PK, Reversade B, Harris RA, Aagaard K, Kristjansson RP, Olafsson I, Eyjolfsson GI, Sigurdardottir O, Iacono WG, Lambalk CB, Montgomery GW, McGue M, Ong KK, Perry JRB, Martin NG, Stefánsson H, Stefánsson K, Boomsma DI (2016). Identification of Common Genetic Variants Influencing Spontaneous Dizygotic Twinning and Female Fertility. *Am J Hum Genet*, 98(5), 898-908. (cites = 64)

Nick Martin's History of the Genetics of Human DZ Twinning

Dorret I. Boomsma

Twin researchers like Nick Martin who love twins for the power of the classical twin design, who work with twin data, recruit twins into their studies, and interact with twins and their relatives become inspired by the data they collect, but also by the questions they are asked by twin families. Nick started the Australian twin register around 1978 from his parental home in Adelaide. He received help from his father, Peter Martin, an influential geneticist who guided him in his first twin study of scholastic ability (published in 1975). His mother, Beryl Martin, a well-known painter of watercolors supplied the 'thank you cards' for twins who registered for the first studies of alcohol metabolism, finger ridge count and, in one of the first discordant-twin design studies, vitamin C and the common cold.

Early on, Nick was confronted by questions from mothers of twins wanting to know 'which set of their twins' he wished to study, why twins seem so plentiful in their family and –most pressing- what the recurrence risk might be of them having another set of twins. It would take nearly 40 years to answer some of their questions. Over these four decades, Nick initiated hormone and ultrasound studies, performed segregation and pedigree analyses, tested candidate genes, carried out linkage projects in sister pairs and took part in large collaborations to enlighten the genetics of dizygotic (DZ) twinning by genome-wide association studies and meta-analysis. Monozygotic twinning was thought be a 'random' event, whereas DZ twins result from the ovulation of two, or more, follicles after processes of follicle growth, selection and ovulation and clearly had a strong familial component.

His first hormone study was published in 1984 and for this project Nick recruited 14 women, eight of whom had at least one set of DZ twins (six with two sets) and six women without DZ twins. Blood draws were at fixed days of their natural cycle, as the hypothesis of interest was that the tendency to have DZ twins is associated with higher follicle-stimulating hormone (FSH) levels in the early follicular phase. Other hormones, including luteinizing hormone (LH) also were assessed. Early FSH and to a lesser extent, LH levels were significantly, at the 10% significance level, higher in mothers of twins.

Martin NG, Olsen ME, Theile H, El Beaini JL, Handelsman D, Bhatnagar AS (1984). Pituitary-ovarian function in mothers who have had two sets of dizygotic twins. *Fertility and Sterility*, 41(6), 878-880. (cites = 53)

In a next paper, published in 1991, Nick increased the number of days for taking blood samples from 3 to 5 and recruited 8 mothers of DZ twins and 8 matched controls and except for one blood draw all blood samples were taken in the women's homes. He set out to test whether multiple ovulation in DZ-twin mothers is because of higher hypothalamic stimulation or whether it is in response to lower serum levels of ovarian inhibin. In this study, FSH levels were again elevated, although not significantly different between the two groups, but LH, inhibin and estradiol levels were elevated on multiple days, suggesting that the primary cause of multiple ovulation is not associated with lower inhibin levels.

Martin NG, Robertson DM, Chenevix-Trench G, de Kretser DM, Osborne J, Burger HG (1991). Elevation of follicular phase inhibin and luteinizing hormone levels in mothers of dizygotic twins suggests nonovarian control of human multiple ovulation. *Fertility and Sterility* 56(3), 469-474. (cites = 39)

Endocrine studies have demonstrated that gonadotrophin release from the hypothalamic-pituitary system is pulsatile. This is observed most strongly for LH release with the signal carried primarily by the frequency of pulses. FSH release is also pulsatile so measuring only mean differences in mean FSH levels may miss effects of changing pulse frequency. Our close collaborator Nils Lambalk, confirmed the earlier endocrine findings for DZ twinning by Nick that women bearing DZ twins have elevated FSH concentration and showed that this increase is particularly associated with a rise in the FSH pulse frequency. These observations were made after serial blood samples were taken every 10 min for for a period of sex hours. Immediately after the last sample, women also received an LHRH challenge.

Lambalk CB, Boomsma DI, Boer L de, Koning CH de, Schoute E, Popp-Snyders C, Schoemaker J (1998). Increased levels and pulsatility of Follicle-Stimulating Hormone in mothers of hereditary dizygotic twins. *J Clinical Endocrinology and Metabolism*, 83, 481-486. (cites = 60)

Nick's 1991 hormone study also looked at whether multiple ovulation occurred in the cycle under study by ovarian ultrasonography on day 12 of the menstrual cycle. Two mothers of DZ twins had more than one follicle. The ultimate ultrasound study was published in the same year, in *Acta Geneticae Medicae et Gemellologiae,* the predecessor of *Twin Research and Human Genetics*, of which Nick became editor in 2000. The year-long study followed 21 mothers of DZ twins and 18 controls and firmly established multiple large follicle growth, both ipsi- and contralateral, in the twin mothers, but not in the controls.

Martin NG, Shanley S, Butt K, Osborne J, O'Brien G (1991). Excessive follicular recruitment and growth in mothers of spontaneous dizygotic twins. *Acta Geneticae Medicae et Gemellologiae,* 40(3-4), 291-301. (cites = 34)

In 1996, the Australian, East-Flanders and Netherlands Twin Register, in collaboration with Cathryn Lewis who then was at the University of Utah in Salt Lake City, attempted to apply complex segregation analysis to determine whether there was evidence for a major gene underlying the tendency to have DZ twins. Large series of pedigrees with mothers of DZ twins as probands were analyzed and the inheritance of DZ twinning was confirmed in these multi-generation data, with paternal as well as maternal inheritance, a low penetrance and limited or no evidence for major gene effects or X-linked inheritance.

Lewis CM, Healey SC, Martin NG (1996). Genetic contribution to DZ twinning. *Am J Medical Genetics*, 61(3), 237-246. (cites = 44)

Meulemans WJ, Lewis CM, Boomsma DI, Derom CA, Berghe H van den, Orlebeke JF, Vlietinck RF, Derom RM (1996). Genetic modeling of dizygotic twinning in pedigrees of spontaneous dizygotic twins. *Am J Medical Genetics*, 61, 258-263. (cites = 50)

Several of Nick's earlier papers mention the strong animal models for DZ twinning, especially the work done in sheep. Maybe naively, at the time we thought that the well-characterized loci leading to multiple ovulation in Booroola merino ewes would hold the key to unraveling multiple ovulation in humans. This led to a long-lasting collaboration with Grant Montgomery, multiple studies of candidate genes and a large linkage study in affected sister pairs. The linkage study, which included affected sister pairs (at least 2 sisters who were both mothers of DZ twins) from 523 families from Australia, New Zealand and some large pedigrees from Utah and the Netherlands did not observe any linkage peaks with LOD scores above 3, and in the

end concluded that "Our data provide further evidence for complex inheritance of familial DZ twinning".

Painter JN, Willemsen G, Nyholt D, Hoekstra C, Duffy DL, Henders AK, Wallace L, Healey S, Cannon-Albright LA, Skolnick M, Martin NG, Boomsma DI, Montgomery GW (2010). A genome wide linkage scan for dizygotic twinning in 525 families of mothers of dizygotic twins. *Hum Reprod*, 25(6), 1569-80. (cites = 34)

The work on candidate genes was summarized by Grant in a review paper published in 2018 and this work also concluded that we still were far away from resolving the genetics on DZ twinning, as rare and low-frequency variants, in e.g. *BMP15* and *BMPR1B*, accounted for only a tiny fraction of variation in DZ twinning.

Gajbhiye R, Fung JN, Montgomery GW (2018). Complex genetics of female fertility. *NPJ Genom Med*, 3, 29.

In the search for common genetic variants that could explain the inheritance of DZ twinning, a new approach became feasible, when genetic variation could be quantified at a large scale using SNP arrays. Hamdi Mbarek, a young scientist working with the Netherlands Twin Register in Amsterdam was courageous enough to analyze the data from 1980 mothers of spontaneous DZ twins and 12,953 controls. These numbers, initially a bit larger, seemed small in comparison to those in increasingly large genome-wide association meta-analyses (GWAMA). And first analyses indeed did not detect much signal in the data. This quite dramatically changed after cleaning the phenotype data and removing ~100 mothers who had conceived their twins after IVF and screening the controls for ties with DZ twinning in their pedigrees. The study identified and replicated, in the DECODE Icelandic databases, an association with DZ twinning for SNPs close to Follicle-Stimulating Hormone Beta Subunit (FSHB) and in SMAD Family Member 3 (SMAD3) providing support for Nick's early studies on FSH concentrations in mothers of DZ twins. The risk alleles close to FSHB and in SMAD3 increased the frequency of twin births in the Icelandic population by 18 and 9%, respectively. FSH always had been a strong candidate for DZ twinning, but the finding for SMAD3 on chromosome 15 was new. The lead SNP associated with DZ twinning maps to the first intron in SMAD3 and is strongly expressed in the human ovary, where it promotes granulosa cell proliferation and steroidogenesis. The region of chromosome 15q22.33 also includes SMAD Family Member 6 (SMAD6). Around the same time, this region was found to be equivalent to genomic region on bovine chromosome 10 (the location of both SMAD3 and SMAD6), which is associated with increasing ovulation rate and twinning in cattle.

Mbarek H, Steinberg S, Nyholt DR, Gordon SD, Miller MB, McRae AF, Hottenga JJ, Day FR, Willemsen G, de Geus EJ, Davies GE, Martin HC, Penninx BW, Jansen R, McAloney K, Vink JM, Kaprio J, Plomin R, Spector TD, Magnusson PK, Reversade B, Harris RA, Aagaard K, Kristjansson RP, Olafsson I, Eyjolfsson GI, Sigurdardottir O, Iacono WG, Lambalk CB, Montgomery GW, McGue M, Ong KK, Perry JRB, Martin NG, Stefánsson H, Stefánsson K, Boomsma DI (2016). Identification of Common Genetic Variants Influencing Spontaneous Dizygotic Twinning and Female Fertility. *Am J Hum Genet*, 98(5), 898-908. (cites = 64)

Efforts to localize and characterize the genes for twinning, and now also for MZ twinning which we for a long time thought to have no genetic basis, continue and are yielding exciting new results. Our last meeting to discuss progress and exciting sets of new results was in

November 2019 in Singapore and promised exciting new insights into the etiology of both types of twinning.

(Twin Workshop Leuven 1998)

(Twin Workshop Leuven 1991)

The Genetics of Biochemical Phenotypes

John B Whitfield

QIMR Berghofer Medical Research Institute, Brisbane

Our collaboration on genetic effects on biochemical characteristics began in 1979 when Nick Martin was at the Australian National University, in Canberra, and exploring the possibility of conducting what became the Alcohol Challenge Twin Study (ACTS) (1). He visited Sydney and came to see me at Royal Prince Alfred Hospital, mainly to ask about laboratory tests to assess subjects' alcohol intake. I had to tell him that the prospects for estimating alcohol intake accurately for an individual person were poor, but we went on to agree that doing a range of biochemical tests on twins and using the results to assess heritability would be valuable. At that time there were few studies of this kind, and they had mostly focused on lipids (particularly cholesterol) because of its relevance to cardiovascular disease.

Over the subsequent forty years our biochemical studies developed through a number of stages, as happened for other phenotypes of biomedical interest. From initial steps to establish the existence of genetic effects and estimate heritability using comparisons of MZ and DZ pair similarity, study size grew to allow consideration of genetic correlations between phenotypes. By about 1990, genotyping of variants in candidate genes was becoming possible and soon after that the typing of genome-wide microsatellite markers led to (mostly unsuccessful) attempts to identify loci affecting quantitative variation by genetic linkage. Around 2005 the technical advances allowing manufacture of genotyping arrays, and the conceptual step from linkage to association testing (2), made genome-wide association studies (GWAS) possible. Possible, that is, if one had access to samples for DNA extraction, phenotypic information, consent from study participants, and funds to purchase the genotyping chips. Fortunately we had the first three and this led gradually to the fourth.

The results of the GWAS revolution are still playing out, but developments so far include not only identification of loci affecting quantitative variation, but a greater understanding of the relationships between phenotypes (including between biomarkers and disease) and increasing use of genetic results to address questions of causation in epidemiology.

Heritability and other twin-pair designs

Blood samples from the ACTS participants were used for a range of biochemical and haematological tests, and the results led to ten papers which tended to estimate heritability and (because a sub-sample of ACTS participants were willing to return for a second time) repeatability. This combination clarified a fact that is still not sufficiently appreciated; when heritability and test-retest repeatability are similar, the long-term average of a diagnostic biomarker or risk factor is strongly dependent on genetic variation and environmental effects tend to be evanescent.

One of these biochemical studies (3) was an early example of integration of a genetic marker into a twin study. It had been known for a long time that serum alkaline phosphatase activity is affected by the ABO and Lewis blood groups, and ABO grouping was one of the tests used to confirm self-reported zygosity in twin-pairs. About 15% of the genetic variance in alkaline phosphatase activity was associated with ABO type – still a large effect even in the GWAS era, and the ABO locus has turned out to be significant (for reasons which are not clear) in GWAS of many phenotypes.

Another variation on twin studies was the use of MZ pairs, and those who participated twice, to assess postulated genetic effects on sensitivity to environmental variation. The hypothesis

(4) was that some variants, which might or might not affect mean values for a phenotype, would affect the response of the phenotype to environmental variation. By genotyping MZ twin pairs for the genetic variant (the MN blood group), and measuring the phenotype (cholesterol) in each twin or in the same person on more than one occasion, it would be possible to test the hypothesis that within-pair or within-person differences would be associated with genotype. Such gene-environment interaction would be of considerable importance if, say, some people obtained benefit from change in diet and others did not. As so often occurs, the original hypothesis was not strongly supported by results (5), but a slightly different one (of effects on triglycerides) emerged. Subsequent multi-centre data (6) suggested that gene-by-environment interaction for lipid levels might exist, with a just-significant result (but for a different locus and phenotype) from genome-wide testing. I mention these studies as an example of an attractive hypothesis, worth some effort to test, not being supported in practice. More generally, GxE interaction has only been shown infrequently despite the large amount of GWAS data now available.

Candidate Genes, Linkage

Association studies involving candidate genes have proved to be a trap, and it is widely accepted that they have led to many false positives through lack of consideration of the multiple testing problem when claiming significant results. The positive aspect has been an increased awareness of the need to set stringent p-values in genome-wide studies and, as far as possible, to replicate results in independent cohorts. Linkage studies for quantitative phenotypes such as biochemical test results have mostly failed for a different reason, because the effect sizes (with few exceptions) are too small to be detectable. Our experience with candidate genes and linkage generally followed this pattern.

One successful candidate gene study was to evaluate the effects of variation at the *HFE* gene, newly found to be necessary (but not sufficient) for haemochromatosis, on serum iron and related measures of iron status in the general population (7). This integrated *HFE* genotype information with the twin study method and showed that although *HFE* variants had significant effects on iron status, they only accounted for a small proportion of the genetic variance – an early example of missing heritability.

Because we had suitable data on related study participants, at first DZ twin pairs and later non-twin siblings, we made a number of attempts to identify loci affecting lipids through linkage analysis but association analyses soon displaced linkage. One successful attempt was for serum butyrylcholinesterase, where a linkage peak was found on chromosome 3, overlapping the *BCHE* gene location. This was later substantiated by GWAS, but it should be admitted that linkage also identified a peak on chromosome 5 which did not show association in the later GWAS.

Blood lead, from h^2 to GWAS

Lead is toxic and widely distributed in the environment, largely because of human mining and industrial processes including previous use in house paints and as a petrol additive. It has been implicated in a range of phenomena from the fall of the Roman empire (now largely refuted, see (8)) to childhood behaviour disorders and educational achievement (for which there is strong evidence of association (9), but many potential confounders which make causation uncertain). Because of the presence of lead in the environment, it was taken for granted that variation in blood lead would be 'environmental' rather than 'genetic'. A series of papers using data from twins and their relatives gave a different perspective.

Firstly, the classical twin method (10) showed evidence for substantial heritability of blood lead concentration in adults ($h^2 \approx 40\%$), with no significant shared environment effect. Linkage

analysis suggested that a region of chromosome 3 contained a variant affecting blood lead. This extensive region includes the *SLC4A7* gene which codes for a transporter affecting lead influx into erythrocytes, which was very encouraging, but this linkage result was not supported by later GWAS results.

Given the evidence for heritability, and the possible localisation of a variant having substantial effects on blood lead, the next step was to conduct a GWAS. This was done in collaboration with Dave Evans and used the ALSPAC cohort from the UK in addition to our data. It found one significant locus (*ALAD*) using a combined sample size of 5400 people. However there was no evidence for significant association at the chromosome 3 linkage region.

This example is important because it follows the stages of genetic investigation from heritability (of a seemingly environmental phenotype) to GWAS, with a diversion through linkage on the way. If it ever becomes possible to gather more data, a much larger GWAS for blood lead should identify more loci and permit the use of Mendelian Randomisation to assess whether associations between lead and childhood development are causal.

GWAS – Heart, Kidney, Liver

The panel of routine diagnostic tests which we ran on blood samples from twins and their families covered a number of organ systems or areas of risk – lipids for heart disease, creatinine, urea and uric acid for kidney function, enzyme tests for liver function, CRP for inflammation.

Although we accumulated biochemical data on around 17,000 adults (mostly with genotyping), the main value of this dataset came from collaboration with other groups who had similar data and from meta-analysis. Through these collaborations, sample sizes in the hundreds of thousands could be achieved and discovery of significant variants has been far beyond what any single group could have managed (11-14). More importantly than listing significant variants, our data contributed to insights such as the causal role of triglycerides in coronary artery disease (15); confirmation that most loci associated with kidney function assessed from creatinine results are also associated with urea and with diagnosed chronic kidney disease (13); and that genes containing variants which affect C-reactive protein concentration cluster in two groups, representing immune and metabolic pathways (12).

GWAS – other phenotypes

Apart from the widely available tests mentioned above, we measured a number of other biochemical phenotypes. Despite the limitations imposed by limited numbers (our studies plus one or just a few others), several important and/or interesting associations have been found.

- As well as blood lead (discussed above), the method for lead estimation also gave results for six other toxic or essential elements in blood cells (As, Cd, Cu, Hg, Se, Zn). These also showed significant heritability, and there was a notable genetic correlation between concentrations of As and Hg ($r_G = 0.83$, whereas $r_E = 0.34$). GWAS for the essential elements, and meta-analysis with similar data from the ALSPAC cohort, showed a number of significant loci for Cu, Se and Zn with, in many cases, probable explanations in terms of gene functions (16).
- An early study on iron and *HFE* genotypes was mentioned above. This was expanded to GWAS with our own data and then to meta-analysis of GWAS data from multiple groups which included almost 50,000 participants. Eleven loci were identified as significant for one or more of the markers of iron status (17), and because of the biological importance of iron and its potential to cause tissue damage there have been a number of attempts to use the relevant genotypes as instrumental variables to test whether associations between iron and disease are causal.

- Plasma cholinesterase (butyrylcholinesterase, BCHE) is an enzyme whose activity is associated with obesity and other aspects of metabolic syndrome, but its function and the reasons for these associations are unknown. Because BCHE measurement was included in our test profile, we carried out a GWAS with the expectation that identification of genes affecting BCHE variation would shed light on its function and relationships with other phenotypes. By far the strongest associations were within or near the *BCHE* gene, and other significant loci were not associated with metabolic risk factors. On the other hand, SNPs in genes associated with metabolic risk tended to have effects on BCHE, suggesting that BCHE variation is a consequence of metabolic abnormalities.
- Carbohydrate-deficient transferrin (CDT) comprises transferrin isoforms which have fewer than the usual four terminal sialic acid residues on their glycan sidechains, and their relative concentration in serum is increased in people with high alcohol intake. Because of our interest in markers of alcohol use, and as an example of variation affecting protein glycosylation, we conducted a GWAS for CDT (18). This identified two loci, the transferrin (*TF*) gene itself and the *PGM1* gene which catalyses an early step in synthesis of the carbohydrate side-chains, showing that variation in both the protein structure and in formation of the glycan component can affect the product.
- Proteolytic cleavage of chromogranins leads to formation of a number of bioactive peptides including catestatin, which has a role in control of blood pressure. Collaboration with Dan O'Connor, the major player in study of chromogranins and related peptides, led us through heritability, linkage and GWAS stages to discovery of two loci affecting catestatin formation (19). Each locus contained a gene for a proteolytic enzyme involved in the intrinsic pathway of coagulation, and review of published literature showed that this process is important for formation of several peptide hormones from their precursors.

Conclusions

Studies on the genetics of biomarkers carry the expectation that because the biomarkers are associated with disease, results will be translatable to the genetics of disease. GWAS results in general may give insight into the mechanisms which regulate or influence the phenotype; they can (depending on the genetic architecture and on study size) predict the phenotype of an individual or stratify their risk of disease; and they can establish or refute causal relationships between apparent risk factors and disease. Genetic studies on biochemical phenotypes have grown and developed over the past forty years, from 412 participants in our early twin studies to over a million in recent collaborative meta-analyses. It should be remembered that the justification for mega-GWAS studies came from initial, smaller GWAS and the justification for the initial GWAS usually came from the knowledge that the phenotypes had significant heritability.

References

1. Martin NG, Perl J, Oakeshott JG, Gibson JB, Starmer GA, Wilks AV. A twin study of ethanol metabolism. BehavGenet 1985;15:93-109.
2. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516-7.
3. Whitfield JB, Martin NG. Determinants of variation in plasma alkaline phosphatase activity: A twin study. AmJHumGenet 1983;35:978-86.

4.   Magnus P, Berg K, Borresen AL, Nance WE. Apparent influence of marker genotypes on variation in serum cholesterol in monozygotic twins. Clin Genet 1981;19:67-70.

5.   Martin NG, Rowell DM, Whitfield JB. Do the mn and jk systems influence environmental variability in serum lipid levels? ClinGenet 1983;24:1-14.

6.   Surakka I, Whitfield JB, Perola M, Visscher PM, Montgomery GW, Falchi M, et al. A genome-wide association study of monozygotic twin-pairs suggests a locus related to variability of serum high-density lipoprotein cholesterol. Twin Res Hum Genet 2012:1-9.

7.   Whitfield JB, Cullen LM, Jazwinska EC, Powell LW, Heath AC, Zhu G, et al. Effects of hfe c282y and h63d polymorphisms and polygenic background on iron stores in a large community sample of twins. AmJHumGenet 2000;66:1246-58.

8.   Retief FP, Cilliers L. Lead poisoning in ancient rome. Acta Theologica 2006;26:147-64.

9.   Bellinger DC. Very low lead exposures and children's neurodevelopment. Curr Opin Pediatr 2008;20:172-7.

10.   Whitfield JB, Dy V, McQuilty R, Zhu G, Heath AC, Montgomery GW, Martin NG. Genetic effects on toxic and essential elements in humans: Arsenic, cadmium, copper, lead, mercury, selenium and zinc in erythrocytes. Environ Health Perspect 2010;118:776-82.

11.   Global Lipids Genetics C, Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45:1274-83.

12.   Ligthart S, Vaez A, Vosa U, Stathopoulou MG, de Vries PS, Prins BP, et al. Genome analyses of >200,000 individuals identify 58 loci for chronic inflammation and highlight pathways that link inflammation and complex disorders. Am J Hum Genet 2018;103:691-706.

13.   Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. Nat Genet 2019;51:957-72.

14.   Tin A, Marten J, Halperin Kuhns VL, Li Y, Wuttke M, Kirsten H, et al. Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. Nat Genet 2019;51:1459-74.

15.   Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. Nat Genet 2013;45:1345-52.

16.   Evans DM, Zhu G, Dy V, Heath AC, Madden PA, Kemp JP, et al. Genome-wide association study identifies loci affecting blood copper, selenium and zinc. Hum Mol Genet 2013.

17.   Benyamin B, Esko T, Ried JS, Radhakrishnan A, Vermeulen SH, Traglia M, et al. Novel loci affecting iron homeostasis and their effects in individuals at risk for hemochromatosis. Nat Commun 2014;5:4926.

18.   Kutalik Z, Benyamin B, Bergmann S, Mooser V, Waeber G, Montgomery GW, et al. Genome-wide association study identifies two loci strongly affecting transferrin glycosylation. Hum Mol Genet 2011;20:3710-7.

19.   Benyamin B, Maihofer AX, Schork AJ, Hamilton BA, Rao F, Schmid-Schonbein GW, et al. Identification of novel loci affecting circulating chromogranins and related peptides. Hum Mol Genet 2017;26:233-42.

Nick Martin and the `Boulder workshops

John K. Hewitt

I first encountered the irrepressible Nick Martin forty-seven years ago. He was working on his doctorate with Lindon Eaves at the University of Birmingham, UK, while I was a newly minted 21 year old graduate of the Psychology department in the process of `jumping ship' by training for a Master's degree in Applied Genetics --- essentially biometrical genetics for those aspiring to careers in animal and plant breeding. That wasn't my intention, as I had already come under the spell of David Fulker's charm, personal generosity, and sharp intellect, and aspired to join the group of behavior geneticists that had formed in Birmingham. Along with David Fulker, these included John Jinks (Chair of Genetics and Lindon Eaves' mentor), Peter Broadhurst (Chair of Psychology and an animal psychogeneticist, the term he preferred), David Hay (at that time working on the genetics of Drosophila behavior) and, of course, Lindon Eaves who was also my mentor for that year of training.

Into this milieu arrived an outgoing and outspoken Australian who was as determined as his mentor was to make the classical twin study a rigorous, innovative, and central research method in human behavior genetics. Like me, Nick Martin was fortunate to find in Lindon Eaves a mentor with the creative genius, mathematical sophistication, and philosophical courage needed to guide the development of his own abilities and research program. Eaves' doctoral thesis, `Aspects of human psychogenetics' (Eaves, 1970) completed not long before, and Nick Martin's own dissertation --- `The classical twin study in human behavior genetics' (Martin, 1976) --- could have served as a prescient call for the methodological developments and training that later developed into the workshops on the `Methodology for genetic studies of twins and families'.

These workshops were first held in Leuven, Belgium, in 1987, 1989, and 1991, organized by Robert Derom and his colleagues under the auspices of the NATO Advanced Study Institutes, and in 1990 and then from 1992 onwards, with funding from the US National Institute on Mental Health (NIMH), annually in Boulder, Colorado. For many years they were known affectionately as the `Twin workshops' and, now that the subject has advanced into the genomic era, simply the `Boulder workshops'. This series is the longest running workshop on behavior genetic methodology, motivating many of its participants to become researchers in behavior genetics (for better or worse).

The workshops have a long history teaching methods and topics at the forefront of the field, introducing new approaches to the genetics community, and inspiring collaborations. The workshop has resulted in the publication of two special editions of the journal Behavior Genetics, the first edited by Nick Martin, Dorret Boomsma, and Mike Neale in 1989, and two textbooks. `Methodology for genetic studies of twins and families' (Neale M.C. and Cardon, 1992) based on the workshops and written by their faculty, became the standard reference for structural equation modeling in behavioral and psychiatric genetics. A second textbook, `Statistical genetics: gene mapping through linkage and association' (Neale B.M., Ferreira, Medland, and Posthuma (Eds.), 2007) addressed the field of genomics that, at that time, was just emerging as a real force in complex trait analysis. To give a sense of the timeliness of the content of the workshops, the second textbook included a chapter (by Patrick Sullivan and Shaun Purcell) on `Analyzing genome-wide association study data: a tutorial using PLINK'. This was at a time when only two GWAS studies, of macular degeneration, had been published,

and prior to the publication of the first ever large scale GWAS (Wellcome Trust Case Control Consortium, 2007).

From the very beginning, and throughout the history of the workshops over the past third of a century, Nick Martin has been perhaps their most enthusiastic supporter and advocate, as well as a very active participant. He is one of the few faculty to have attended every single workshop to date, and he has always been ready to present the latest developments from his research group (often on a Friday afternoon when it has now become a tradition to wrap up the workshop with short overviews of what research most excites the faculty). Perhaps more importantly, he has shaped the workshops with his strong convictions about what the workshops' goals for the students should be; he is, perhaps as much as any of the faculty, firmly committed to their didactic mission. Nick has always found a way to bring his students to the workshops, and his reward has been watching them advance from students to faculty to directors of the workshop --- Sarah Medland and Dave Evans are two of Nick's former students who currently share the workshop academic leadership with Mike Neale and Ben Neale.

In Nick's view, if one can presume to speak for him, the first goal of the workshop should be to instill a clear sense of what the central methods of biometrical behavior genetics can tell us. No one should leave the workshop after the first day without this basic understanding. This is essential, whether achieved through the straightforward ACE model analysis of the classical twin study, or through tracing the historical development of the subject from Mendel, Galton, and Fisher through to genome wide association studies and sequencing. Nick has always been an enthusiast for summarizing the first day's work on the whiteboard, prompting students to share their results and hazard their interpretations of what they might mean. Technicalities are important of course, but not at the expense of grasping what the results are telling us --- is the trait under study influenced by genetics, by shared family environments, a lot, a little, or not at all? Why? How can we know this?

The second goal of the workshop should be to present a state-of-the-art version of what was one of the most significant advances in the field of behavior genetics, developed in Nick's doctoral thesis and published in the seminal paper by Martin NG and Eaves LJ (1977) The genetical analysis of covariance structure. Heredity 38(1):79–95. Multivariate genetic and environmental analysis was, arguably, one of the most consequential advances in our field, leading to the wide application of structural equation modeling for twin and family study data and, most recently, genomic data.

Whenever we have been tempted to crowd out either of these two goals to make room for the myriad technical advances in our subject, Nick has reminded us that these central themes form the basis of our subject and our workshops should emphasize them just as much as the rapidly occurring and powerful new developments.

Alongside the academic goals of the workshops, Nick firmly believes in the value of the social interactions that the workshops encourage in a way that few other learning experiences can. This is how scientific collaborations and personal relationships can get their start, and new ideas, research projects, and publications can follow. To facilitate this, Nick is the ready and willing cheerleader for faculty and student introductions on the first day, emphasizing the remarkably international and, indeed, multinational background of many of the workshop faculty, and calling on the students who work in similar research areas to identify themselves and recognize their potential new colleagues. It helps to be an extravert and to have an energy level that is still phenomenal even as Prof. Martin celebrates his 70th birthday. After a full day

traveling from Brisbane to Boulder, many of us would just want to crash out in our hotel room. But one of the attractions of the Boulder is that there's a ski resort an hour away, and that's where you'll likely find Nick in the time available between his arrival in Boulder and the first meeting of the workshop faculty to finalize the program for the coming week.

Happy 70th Birthday, Nick, from all of us at the Boulder workshop!

John K. Hewitt, 02-19-2020, Boulder, Colorado.

(Twin Workshop Boulder 1990)

Nick Martin and The Extended Twin Model

Hermine H. Maes

I first met Nick at the very first 'Twin Methodology' workshop held in Leuven, Belgium in 1987, and I've had the pleasure of seeing him at almost every workshop since (at 34 right now and counting). I was lucky enough to be able to attend the first workshop as it was held at my alma mater, then helped organize the next and from then on got invited to help teach them. Nick and I have spent countless sessions teaching 'the ACE model 101' together to hundreds of workshop participants, with the same level of enthusiasm from Nick as at that first workshop. It was this enthusiasm for science and the desire to improve how investigators analyze their data that attracted me to pursue this line of research and led me to move to Richmond to do a postdoc with Lindon Eaves.

Although Nick had already moved on from Richmond to Brisbane to start his own 'Genetic Epidemiology' unit, his first major data collection project was clearly inspired by the work he had done with Lindon. They had conceived the 'extended twin model'. Recognizing the limitations of the classical twin study, which typically partitions the variance in a trait into additive genetic (A), common (C) and unique (E) environmental sources, they sought to extend it to include other relatives - parents of twins, siblings of twins, spouses of twins and children of twins. This extension allows one to not only evaluate the consistency of the estimates of the genetic and environmental contributions to the variance across a range of relationships, but also estimate additional sources of variance previously confounded with other estimates. The extended twin model (Eaves LJ et al. 1999) provides a test for environmental or cultural transmission besides genetic transmission thus dividing the shared environment into sources shared with parents and those shared with siblings but not parents. In addition, dominance variance can now be simultaneously estimated with shared environmental variance. Further excess environmental sharing in twins compared to siblings can be quantified as 'special twin environment' or reflect potential age-specific effects of genes. Sex differences in all these sources of variance can equally be evaluated. Finally, relationship through marriage provides information about the extent of assortative mating.

Thus from the theory of the causes of variation in human behavior, they developed a model system (Truett KR et al. 1994) for the analysis of family resemblance in extended kinships of twins, and collected data on health and lifestyle from a large sample of twins and their relatives, then fitted their model to the data and started the 'stealth' revolution. A path diagram of the model resembled a stealth bomber, a fitting name for a powerful model. Questionnaire data were collected on thousands of twins and their relatives both in Virginia (the Virginia 30,000) and in Australia (the Australia 25,000), allowing researchers to this day to explore the complexities of the causes of variation in complex traits ranging from social attitudes (Eaves L et al. 1999), depressive symptoms (Kendler KS et al. 1994), panic and phobias (Kendler KS et al. 1995), body mass index (Maes HH et al. 1997, Bergin JE et al. 2012), church attendance (Kirk KM et al. 1999), alcohol use (Maes HH et al. 1999, Verhulst B et al. 2018), neuroticism (Lake RI et al. 2000, Boomsma DI et al. 2018), smoking initiation (Maes HH et al. 2006, Maes HH et al. 2018), political attitudes (Hatemi PK et al. 2009) and so on. Many of these publications would not have happened, was it not for Nick's generosity of data, time, encouragement, travel assistance and hospitality, discussing results over wine and good food, often accompanied by excellent classical music.

Through these interactions, Nick inspired graduate students and postdocs to further explore extensions of the extended twin (ET) model. While the first iterations and applications of the ET model were written in Fortran, we developed code in Mx and later in OpenMx (Maes HH et al. 2009) that allowed fitting it to raw data, applying it to continuous and categorical data, incorporating covariates and extending it to the multivariate case (Maes HH, Neale MC 1999). Alternative mechanisms of intergenerational transmission and assortment - phenotypic cultural transmission and social homogamy - were coded (Keller MC et al. 2009). New programs were written to simulate data to evaluate bias, precision and accuracy of the parameter estimates (Coventry WL et al. 2005, Keller MC et al. 2010), as well as power associated with different family structures (Medland SE et al. 2009). Additional relatives (Vinkhuyzen AA et al. 2012) and non-biologically related family members (Maes HH et al. 2007, Leve LD et al. 2018) were included. The utility of subsets of the data, such as children of twins (COT), to disentangle genetic from cultural transmission were explored (Eaves LJ et al. 2005, Docherty AR et al. 2015) and expanded (Marceau K et al. 2015, McAdams TA et al. 2018). Cross-cultural comparisons were undertaken to test the reproducibility and consistency of findings (Lake RI, Eaves LJ 2000, Maes HH, Morley K 2018). Twin registries were expanded with data collected from other relatives (Kaprio J et al. 1987, Boomsma DI et al. 2008, Ligthart L et al. 2019), recognizing the added value, not just in terms of power but in capturing more of the nuances of how genetic and environmental factors act and interact in creating individual differences. The list of phenotypes to which these models have been applied continues to grow, with publications on brain structure (Posthuma D et al. 2000), blood pressure (Kupper N et al. 2005), parturition timing (Kistka ZA et al. 2008), personality disorder (Distel MA et al. 2009), intelligence (Vinkhuyzen AA, van der Sluis S 2012), political orientation (Kandler C et al. 2012), pro-inflammatory state (Neijts M et al. 2013), personality (Hahn E et al. 2012, Kandler C et al. 2019), and political affiliation (Kornadt AE et al. 2018, Hufer A et al. 2019). On a personal note, Nick has been extremely supportive in my career - and deserves every spot as co-author and contributor. Furthermore, he genuinely cares about moving the field of (behavior) genetics forward, and has clearly put his stamp on developing models for extended twin kinships, collecting relevant data and fitting ET models to them, and through it all mentored and encouraged his academic extended family, while enjoying their company during 'just bring food' dinners, good wine, and if possible listening some lovely music.

Thanks Nick!

References

1. Eaves LJ, et al, *Biological and Cultural Inheritance of Stature and Attitudes*, in *Personality and Psychopathology*, Cloninger CR, Editor. 1999, American Psychiatric Press, Inc: Washington, DC. p. 269-308.
2. Truett KR, et al (1994) A model system for analysis of family resemblance in extended kinships of twins. Behav Genet, 24(1): 35-49.
3. Eaves L, et al (1999) Comparing the biological and cultural inheritance of personality and social attitudes in the Virginia 30,000 study of twins and their relatives. Twin Res, 2(2): 62-80.
4. Kendler KS, et al (1994) Sources of individual differences in depressive symptoms: analysis of two samples of twins and their families. Am J Psychiatry, 151(11): 1605-14.
5. Kendler KS, et al (1995) A twin-family study of self-report symptoms of panic-phobia and somatization. Behav Genet, 25(6): 499-515.
6. Maes HH, et al (1997) Genetic and environmental factors in relative body weight and human adiposity. Behav Genet, 27(4): 325-51.

7. Bergin JE, et al (2012) Genetic and environmental transmission of body mass index fluctuation. Behav Genet, 42(6): 867-74.
8. Kirk KM, et al (1999) Frequency of church attendance in Australia and the United States: models of family resemblance. Twin Res, 2(2): 99-107.
9. Maes HH, et al (1999) Religious attendance and frequency of alcohol use: same genes or same environments: a bivariate extended twin kinship model. Twin Res, 2(2): 169-79.
10. Verhulst B, et al (2018) Extended Twin Study of Alcohol Use in Virginia and Australia. Twin Res Hum Genet, 21(3): 163-178.
11. Lake RI, et al (2000) Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. Behav Genet, 30(3): 223-33.
12. Boomsma DI, et al (2018) An Extended Twin-Pedigree Study of Neuroticism in the Netherlands Twin Register. Behav Genet, 48(1): 1-11.
13. Maes HH, et al (2006) Genetic and cultural transmission of smoking initiation: an extended twin kinship model. Behav Genet, 36(6): 795-808.
14. Maes HH, et al (2018) Cross-Cultural Comparison of Genetic and Cultural Transmission of Smoking Initiation Using an Extended Twin Kinship Model. Twin Res Hum Genet, 21(3): 179-190.
15. Hatemi PK, et al (2009) Genetic and environmental transmission of political attitudes over a life time. J Politics, 71: 1141-1156.
16. Maes HH, et al (2009) Flexible Mx specification of various extended twin kinship designs. Twin Res Hum Genet, 12(1): 26-34.
17. Keller MC, et al (2009) Modeling extended twin family data I: description of the Cascade model. Twin Res Hum Genet, 12(1): 8-18.
18. Coventry WL, Keller MC (2005) Estimating the extent of parameter bias in the classical twin design: a comparison of parameter estimates from extended twin-family and classical twin designs. Twin Res Hum Genet, 8(3): 214-23.
19. Keller MC, et al (2010) Are extended twin family designs worth the trouble? A comparison of the bias, precision, and accuracy of parameters estimated in four twin family models. Behav Genet, 40(3): 377-93.
20. Medland SE, Keller MC (2009) Modeling extended twin family data II: power associated with different family structures. Twin Res Hum Genet, 12(1): 19-25.
21. Vinkhuyzen AA, et al (2012) Reconsidering the heritability of intelligence in adulthood: taking assortative mating and cultural transmission into account. Behav Genet, 42(2): 187-98.
22. Maes HH, et al (2007) Genetic and cultural transmission of antisocial behavior: an extended twin parent model. Twin Res Hum Genet, 10(1): 136-50.
23. Leve LD, et al (2018) Naturalistic Experimental Designs as Tools for Understanding the Role of Genes and the Environment in Prevention Research. Prev Sci, 19(1): 68-78.
24. Eaves LJ, et al (2005) Revisiting the children of twins: can they be used to resolve the environmental effects of dyadic parental treatment on child behavior? Twin Res Hum Genet, 8(4): 283-90.
25. Docherty AR, et al (2015) Comparison of Twin and Extended Pedigree Designs for Obtaining Heritability Estimates. Behav Genet, 45(4): 461-6.
26. Marceau K, et al (2015) Parental knowledge is an environmental influence on adolescent externalizing. J Child Psychol Psychiatry, 56(2): 130-7.
27. McAdams TA, et al (2018) Revisiting the Children-of-Twins Design: Improving Existing Models for the Exploration of Intergenerational Associations. Behav Genet, 48: 397-412.

28. Kaprio J, et al (1987) Design and sampling considerations, response rates, and representativeness in a Finnish Twin Family Study. Acta Genet Med Gemellol (Roma), 36(1): 79-93.
29. Boomsma DI, et al (2008) Design and implementation of a twin-family database for behavior genetics and genomics studies. Twin Res Hum Genet, 11(3): 342-8.
30. Ligthart L, et al (2019) The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-Family Designs. Twin Res Hum Genet: 1-14.
31. Posthuma D, et al (2000) Multivariate genetic analysis of brain structure in an extended twin design. Behav Genet, 30(4): 311-9.
32. Kupper N, et al (2005) Heritability of daytime ambulatory blood pressure in an extended twin design. Hypertension, 45(1): 80-5.
33. Kistka ZA, et al (2008) Heritability of parturition timing: an extended twin design analysis. Am J Obstet Gynecol, 199(1): 43 e1-5.
34. Distel MA, et al (2009) Familial resemblance of borderline personality disorder features: genetic or cultural transmission? PLoS One, 4(4): e5334.
35. Kandler C, et al (2012) Left or right? Sources of political orientation: the roles of genetic factors, cultural transmission, assortative mating, and personality. J Pers Soc Psychol, 102(3): 633-45.
36. Neijts M, et al (2013) Genetic architecture of the pro-inflammatory state in an extended twin-family design. Twin Res Hum Genet, 16(5): 931-40.
37. Hahn E, et al (2012) The complexity of personality: advantages of a genetically sensitive multi-group design. Behav Genet, 42(2): 221-33.
38. Kandler C, et al (2019) The Study of Personality Architecture and Dynamics (SPeADy): A Longitudinal and Extended Twin Family Study. Twin Res Hum Genet: 1-6.
39. Kornadt AE, et al (2018) On the genetic and environmental sources of social and political participation in adolescence and early adulthood. PLoS One, 13(8): e0202518.
40. Hufer A, et al (2019) Genetic and environmental variation in political orientation in adolescence and early adulthood: A Nuclear Twin Family analysis. J Pers Soc Psychol.

(Christmas in Crete)

Eaves LJ, Last KA, Young PA, Martin NG (1978) Model-fitting approaches to the analysis of human behavior. Heredity 41, 249-320 (cites = 498)

Martin NG, Eaves LJ, Kearsey MJ, Davies P (1978) The power of the classical twin study. Heredity 40, 97-116. (cites = 501)

Martin NG, Eaves LJ (1977) The genetical analysis of covariance structure. Heredity, 38, 79-95. (cites = 436)

Statistical Power and the Classical Twin Design

Sham PC, Purcell SP, Cherny SS, Neale MC, Neale BM

Dr Nick Martin is one of the most prolific and influential behavioral geneticists in the world, who has also been a key motivator, teacher and role model for his students, including ourselves. Over the years, we have greatly benefitted from Nick's wonderful teaching, very often demonstrating how theory can be applied in practice to investigate interesting and important scientific questions, and providing a much-needed historical perspective on the latest developments in our fast-moving field. It is therefore our great honour and privilege to review one of Nick's earliest papers, in celebration of his 70th birthday.

The paper "The power of the classical twin study" (Martin et al, 1978) was based on work from Nick's PhD thesis (Martin, 1976), completed in the Department of Genetics at the University of Birmingham. It was in this department that the field of biometrical genetics (Mather & Jinks, 1982; Evans et al, 2002) was established by pioneers who included Kenneth Mather, John Jinks, David Fulker and Lindon Eaves (Nick's PhD supervisor). The principles of biometrical genetics, as compared to other contemporary approaches to the analysis family data, were laid down in a seminal paper from the department (Jinks & Fulker, 1970).

While the aim of biometrical genetics was to partition the sources of individual differences in the population according to various genetic and environmental sources of variation, Jinks and Fulker recognised that the ability to untangle different sources of variation from one another requires certain minimal experimental conditions – the "minimum data". For example, an analysis of variance for monozygotic twins reared apart would yield 2 summary statistics (the between-group mean squares and the within-group mean-squares) which, when equated to the theoretical expected mean squares under the classical quantitative genetic model, would provide estimates for the total genetic and the total environmental variances, but would not be able to separate out additive effects from dominance, nor the familial environment shared by siblings reared together from environmental influences unique to each sibling. A study that includes more varieties of relationships would provide more summary statistics, which would enable more sources of variation to be separately estimated.

Martin et al (1978) recognized that, even when an experimental design would provide the "minimum data" for resolving certain sources of variation, the probability of achieving this in practice would still depend on having a sufficient sample size – "If the power of a study to detect a given effect is low and in fact we do not find evidence for the effect in our sample then we should be foolish to infer that the effect is not present in the population". They pointed out that theoretical power calculations in the literature at the time dealt with "human experimental designs which are seldom (if ever) used", but not "the classical twin design, the most common design in human biometrical genetics".

The paper then went on to describe an analytical approach to perform a power calculation for the classical twin design. The method involved calculating the expected values of the observed mean squares under the specified parameter values of a true model, and then equating these to the theoretical expected mean squares under a false model to estimate the parameters of the false model (using iterative weighted least squares). By substituting the expected mean squares under the true model as the observed mean squares of a goodness-of-fit chi-square test statistic for the false model, they obtained the non-centrality parameter of the distribution of the test statistic. This enabled them to calculate the approximate power of the test for any desired

significance level. Because the non-centrality parameter is proportional to sample size, the results can be easily extrapolated to calculate the power for any sample size, and to calculate the required sample size for any desired power. The accuracy of the power estimates obtained from the non-central chi-squared distribution were shown to be acceptable by simulation, for a range of parameter values and sample sizes. Using this method, it was shown that 600 twin pairs were required to reject most false models, and that an optimal proportion of monozygotic (MZ) and dizygotic (DZ) twin pairs under most true models was between 1/3 and 1/2. The paper ended with a section on the power of detecting non-additive and directional effects, with 3 subsections: (i) GxE interaction, by regressing pair variances on pair means, (ii) directional dominance, by testing the phenotypic distribution for skewness, and (iii) directional allele frequency differences, again by testing the phenotypic distribution for skewness.

Two other papers from Nick and colleagues published at around the same time (Martin & Eaves, 1977; Eaves et al, 1978) were extremely influential in clarifying the properties of existing analytic approaches to family data that use raw data, correlations, or mean squares as the starting point. They also introduced the use of covariance matrices as an alternative, and integrated factor analysis methodology into biometrical genetic analysis. These two papers, together with Martin et al (1978), laid much of the foundation for the later developments in human behavior genetics, including the establishment of large twin registries and the development of modern maximum likelihood approaches for model estimation and testing that enabled the extension of the classical twin model to threshold traits, multiple phenotypes, and extended twin-families (Neale & Cardon, 1992).

Power calculation has remained an important issue in human genetics research. Subsequent papers to Martin et al (1978) have considered the power of new study designs including threshold traits (Neale et al, 1994), multivariate phenotypes (Schmitz et al, 1998), and extended twin designs (Posthuma & Boomsma, 2000). As the field moved to include molecular data for gene mapping, analytic power calculations were developed for quantitative trait linkage and association analyses under the variance components model, also using the non-central chi-squared distribution (Nance & Neale 1989; Sham et al, 2000; Purcell et al, 2003). In the genome-wide association studies (GWAS) era, the variance components model has been applied to estimate the heritability attributable to common single nucleotide polymorphisms (SNPs), and the power of this approach has also be characterised (Visscher et al, 2014).

The seminal paper of Martin et al (1978) on the power of the classical twin design was revisited by Visscher (2004) who calculated power via the standard errors of the variance components and the expected values of the maximum likelihood ratio test statistics. His results are largely comparable to those of Martin et al (1978), with the major difference being that the consideration of likelihood ratio statistics enabled a specific parameter in a model to be tested (e.g., the additive genetic effects within a full model that also contains shared sibship environment, and individual-specific environment), rather than the entire model.

By highlighting statistical power considerations, Nick calls to mind Ronald Fisher, who in his Presidential Address to the First Indian Statistical Congress said "to consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of." The power calculations Martin and colleagues are exactly the kind of prospective treatments that Fisher no doubt had in mind when he quipped as such.

As a pioneer of the fields of biometrical and behavioral genetics, Nick's knowledge, insights and perspective have benefitted entire generations of reseachers in behavioral genetics who have attended the annual "twin workshops", often multiple times. We were fortunate to progress to faculty members of the workshop, and have more directly experienced Nick's enthusiasm and intellectual curiosity, greatly facilitating the sharing of ideas and lively debates, not only among faculty members, but also with the students. These debates and discussions were what have made the workshops so enjoyable, and often led to new and fruitful research directions. On the occasion of Nick's 70th birthday, we express our appreciation and gratitude to him, glance backward to what we have achieved, and look forward to working together to extend the frontiers of the field.

References

Eaves LJ, Last KA, Young PA, Martin NG (1978) Model-fitting approaches to the analysis of human behavior. *Heredity* 41, 249-320

Evans DM, Gillespie NA, Martin NG (2002) Biometrical genetics. *Biological Psychology* 61, 33-51

Fisher RA (1938). Presidential address to the first Indian statistical congress. *Sankhya*, 4, 14-17.

Jinks JL, Fulker DW (1970) Comparison of the biometrical, genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychological Bulletin* 73, 311-349

Mather K, Jinks JL (1982) Biometrical Genetics. Chapman & Hall, New York.

Martin NG (1976) The classical twin study in human behavioural genetics. PhD thesis. University of Birmingham

Martin NG, Eaves LJ (1977) The genetical analysis of covariance structure. *Heredity*, 38, 79-95

Martin NG, Eaves LJ, Kearsey MJ, Davies P (1978) The power of the classical twin study. *Heredity* 40, 97-116

Nance WE, Neale MC (1989). Partitioned twin analysis: A power study. Behavior Genetics, 19, 143-150.

Neale MC, Cardon LR (1992) Methodology for genetic studies of twins and families. Kluwer Academic Publishers, Dordrecht.

Neale MC, Eaves LJ, Kendler KS (1994) The power of the classical twin study to resolve variation in threshold traits. *Behavior Genetics* 24. 239-258

Posthuma D, Boomsma DI (2000) A note on the statistical power in extended twin designs. *Behavior Genetics* 30, 147-158

Schmitz S, Cherny SS, Fulker DW (1998) Increase in power through multivariate analyses. *Behavior Genetics* 28, 357-363

Sham PC, Cherny SS, Purcell S, & Hewitt JK (2000) Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. *American Journal of Human Genetics*, 66, 1616-1630.

Purcell S, Cherny SS, Sham PC (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19, 149-150.

Visscher PM (2004) Power of the classical twin design revisited. *Twin Research* 7, 505-512

Visccher PM, Hemani G, Vinkhuyzen AAE, Chen GB, Lee AH, Qray NR, Goddard ME, Yang J (2014) Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLOS Genetics* 10, e1004269

Gene Discovery Using Twins

David Duffy, Rick Sturm, Gu Zhu, Stuart Macgregor

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 honour of having the highest incidence of the disease in the world (swapping occasionally with New Zealand). The Queensland Melanoma Project (at PAH) 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 naevi (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 colour 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 modelling the genetics of multivariate and time series data - we were just moving to use of the structural equation modelling 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 wouldn't 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) too 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 digital image all the moles on a subset of the BTNS that 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 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 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 (G. Zhu et al., 1999), we reported on results from candidate microsatellite (and even a couple of SNP) marker genotyping of 352 families (a number that now seems hilariously small), where we find a quantitative trait locus close to the familial melanoma gene *CDKN2A* explained 27% of total variance in total body mole count (lod 2.6, $P=6\times10^{-4}$). We concluded at the time that the actual naevus 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 GWAS meta-analyses, we confirm 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 colour, with the advent of digital photography in the early 2000's greatly enhancing the phenotypic characteristics of the iris that could be captured and studied in the twins. It had been known that blue eye colour 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 (Gu 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 colour. By 2007 (Duffy et al., 2007), we had identified a three-SNP haplotype that almost completely explained blue eye colour, that lay not in *OCA2* itself, but in an intron of the neighbouring *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 eminent in the study of human pigmentation genetics and the melanocortin-1 receptor (gene *MC1R*) in particular. Valverde and coworkers (Valverde et al., 1995) had reported variants in MC1R were associated with fair skin and red hair, and shown 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 (Palmer et al., 2000), we showed that risk of melanoma due to carrying *MC1R* red-hair variants was not completely explainable by measured skin colour. This finding has been extensively replicated, and it is clear that the effects of *MC1R* are not just on colouring (changing the melanin in the skin from dark eumelanin to a mixture of eumelanin and red-brown pheomelanin), but also on 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).
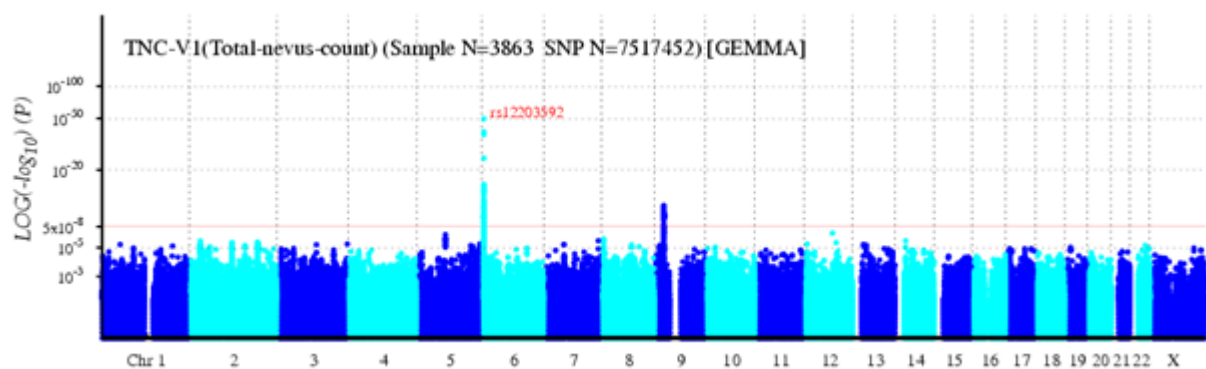

Figure 1. GWAS for total nevus count in 12 year-old twins from BTNS as of January 2020 (from analysis by Gu Zhu). The two most significant associations are over the *IRF4* (chr 6) and *MTAP* (chr 9) genes, but *PLA2G6* on chr 22 still doesn't get a look in – further evidence of age heterogeneity.

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 (eg *R/r)*, but fall in *R/R* homozygotes. In the case of the rs12203592 polymorphism in *IRF4*, associated with skin and hair colour by us and others in a multi-country consortium analysis (Han et al., 2008; Sulem et al., 2007), we had seen association to a nearby SNP, but the strength of the association to rs12203592 and in 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 one looked at raised moles or a flat moles in the BTNS twins. There allele associated with high total mole count at this locus also changed depending on age, when we compared the young twins to either their own parents, or to other populations of adults.

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 52000 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 controlling mole count and melanoma risk, most affecting both traits equally, but some just for mole count (for example, *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. We 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 colour 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. Whilst the first of these studies employed the linkage approach, the strategy really began to yield genes when the genome-wide association study (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 which 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)(Hysi et al., 2010), 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 mean 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 which were subsequently shown to influence melanoma risk. In the case of eye disease, the same genes which influence a person's risk of intraocular pressure in early/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 endophenotype 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 very 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 endeavours.

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*(4), 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*(11), 1891–1897.

Craig, J. E., Han, X., Qassim, A., Hassall, M., Cooke Bailey, J. N., Kinzy, T. G., Khawaja, A. P., An, J., Marshall, H., Gharahkhani, P., Igo, R. P., Jr, Graham, S. L., Healey, P. R., Ong, J.-S., Zhou, T., Siggs, O., Law, M. H., Souzeau, E., Ridge, B., … MacGregor, S. (2020). Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. *Nature Genetics*. https://doi.org/10.1038/s41588-019-0556-y

Duffy, D. L., Iles, M. M., Glass, D., Zhu, G., Barrett, J. H., Höiom, V., Zhao, Z. Z., Sturm, R. A., Soranzo, N., Hammond, C., Kvaskoff, M., Whiteman, D. C., Mangino, M., Hansson, J., Newton-Bishop, J. A., GenoMEL, Bataille, V., Hayward, N. K., Martin, N. G., … Montgomery, G. W. (2010). IRF4 variants have age-specific effects on nevus count and predispose to melanoma. *American Journal of Human Genetics*, *87*(1), 6–16.

Duffy, D. L., Montgomery, G. W., Chen, W., Zhao, Z. Z., Le, L., James, M. R., Hayward, N. K., Martin, N. G., & 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*(2), 241–252.

Duffy, D. L., Zhu, G., Li, X., Sanna, M., Iles, M. M., Jacobs, L. C., Evans, D. M., Yazar, S., Beesley, J., Law, M. H., Kraft, P., Visconti, A., Taylor, J. C., Liu, F., Wright, M. J., Henders, A. K., Bowdler, L., Glass, D., Ikram, M. A., … Martin, N. G. (2018). Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. *Nature Communications*, *9*(1), 4774.

Falchi, M., Bataille, V., Hayward, N. K., Duffy, D. L., Bishop, J. A. N., Pastinen, T., Cervino, A., Zhao, Z. Z., Deloukas, P., Soranzo, N., Elder, D. E., Barrett, J. H., Martin, N. G., Bishop, D. T., Montgomery, G. W., & Spector, T. D. (2009). Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. *Nature Genetics*, *41*(8), 915–919.

Han, J., Kraft, P., Nan, H., Guo, Q., Chen, C., Qureshi, A., Hankinson, S. E., Hu, F. B., Duffy, D. L., Zhao, Z. Z., Martin, N. G., Montgomery, G. W., Hayward, N. K., Thomas, G., Hoover, R. N., Chanock, S., & Hunter, D. J. (2008). A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genetics*, *4*(5), e1000074.

Hysi, P. G., Cheng, C.-Y., Springelkamp, H., Macgregor, S., Bailey, J. N. C., Wojciechowski, R., Vitart, V., Nag, A., Hewitt, A. W., Höhn, R., Venturini, C., Mirshahi, A., Ramdas, W. D., Thorleifsson, G., Vithana, E., Khor, C.-C., Stefansson, A. B., Liao, J., Haines, J. L., … Aung, T. (2014). Genome-wide analysis of multi-ancestry cohorts identifies new loci influencing intraocular pressure and susceptibility to glaucoma. *Nature Genetics*, *46*(10), 1126–1130.

Hysi, P. G., Young, T. L., Mackey, D. A., Andrew, T., Fernández-Medarde, A., Solouki, A. M., Hewitt, A. W., Macgregor, S., Vingerling, J. R., Li, Y.-J., Ikram, M. K., Fai, L. Y., Sham, P. C., Manyes, L., Porteros, A., Lopes, M. C., Carbonaro, F., Fahy, S. J., Martin, N. G., … Hammond, C. J. (2010). A genome-wide association study for myopia and refractive error identifies a susceptibility locus at 15q25. *Nature Genetics*, *42*(10), 902–905.

Ioannidis, J. P. A., Klavans, R., & Boyack, K. W. (2018). Thousands of scientists publish a paper every five days. *Nature*, *561*(7722), 167–169.

Law, M. H., Bishop, D. T., Lee, J. E., Brossard, M., Martin, N. G., Moses, E. K., Song, F., Barrett, J. H., Kumar, R., Easton, D. F., Pharoah, P. D. P., Swerdlow, A. J., Kypreou, K. P., Taylor, J. C., Harland, M., Randerson-Moor, J., Akslen, L. A., Andresen, P. A., Avril, M.-F., … Iles, M. M. (2015). Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. *Nature Genetics*, *47*(9), 987–995.

Lu, Y., Vitart, V., Burdon, K. P., Khor, C. C., Bykhovskaya, Y., Mirshahi, A., Hewitt, A. W., Koehn, D., Hysi, P. G., Ramdas, W. D., Zeller, T., Vithana, E. N., Cornes, B. K., Tay, W.-T., Tai, E. S., Cheng, C.-Y., Liu, J., Foo, J.-N., Saw, S. M., … Wong, T. Y. (2013). Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus. *Nature Genetics*, *45*(2), 155–163.

Macgregor, S., Hewitt, A. W., Hysi, P. G., Ruddle, J. B., Medland, S. E., Henders, A. K., Gordon, S. D., Andrew, T., McEvoy, B., Sanfilippo, P. G., Carbonaro, F., Tah, V., Li, Y. J., Bennett, S. L., Craig, J. E., Montgomery, G. W., Tran-Viet, K.-N., Brown, N. L., Spector, T. D., … Mackey, D. A. (2010). Genome-wide association identifies ATOH7 as a major gene determining human optic disc size. *Human Molecular Genetics*, *19*(13), 2716–2724.

MacGregor, S., Ong, J.-S., An, J., Han, X., Zhou, T., Siggs, O. M., Law, M. H., Souzeau, E., Sharma, S., Lynn, D. J., Beesley, J., Sheldrick, B., Mills, R. A., Landers, J., Ruddle, J. B., Graham, S. L., Healey, P. R., White, A. J. R., Casson, R. J., … Hewitt, A. W. (2018). Genome-wide association study of intraocular pressure uncovers new pathways to glaucoma. *Nature Genetics*, *50*(8), 1067–1071.

McGregor, B., Pfitzner, J., Zhu, G., Grace, M., Eldridge, A., Pearson, J., Mayne, C., Aitken, J. F., Green, A. C., & Martin, N. G. (1999). Genetic and environmental contributions to size, color, shape, and other characteristics of melanocytic naevi in a sample of adolescent twins. In *Genetic Epidemiology* (Vol. 16, Issue 1, pp. 40–53). https://doi.org/3.0.co;2-1">10.1002/(sici)1098-2272(1999)16:1<40::aid-gepi4>3.0.co;2-1

Palmer, J. S., Duffy, D. L., Box, N. F., Aitken, J. F., O'Gorman, L. E., Green, A. C., Hayward, N. K., Martin, N. G., & 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*(1), 176–186.

Powell, J. E., Henders, A. K., McRae, A. F., Caracella, A., Smith, S., Wright, M. J., Whitfield, J. B., Dermitzakis, E. T., Martin, N. G., Visscher, P. M., & Montgomery, G. W. (2012). The Brisbane Systems Genetics Study: genetical genomics meets complex trait genetics. *PloS One*, *7*(4), e35430.

Praetorius, C., Grill, C., Stacey, S. N., Metcalf, A. M., Gorkin, D. U., Robinson, K. C., Van Otterloo, E., Kim, R. S. Q., Bergsteinsdottir, K., Ogmundsdottir, M. H., Magnusdottir, E., Mishra, P. J., Davis, S. R., Guo, T., Zaidi, M. R., Helgason, A. S., Sigurdsson, M. I., Meltzer, P. S., Merlino, G., … Steingrimsson, E. (2013). A polymorphism in IRF4 affects human pigmentation through a tyrosinase-dependent MITF/TFAP2A pathway. *Cell*, *155*(5), 1022–1033.

Sturm, R. A., Duffy, D. L., Box, N. F., Chen, W., Smit, D. J., Brown, D. L., Stow, J. L., Leonard, J. H., & Martin, N. G. (2003). The role of melanocortin-1 receptor polymorphism in skin cancer risk phenotypes. *Pigment Cell Research / Sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society*, *16*(3), 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*(2), 424–431.

Sulem, P., Gudbjartsson, D. F., Stacey, S. N., Helgason, A., Rafnar, T., Magnusson, K. P., Manolescu, A., Karason, A., Palsson, A., Thorleifsson, G., Jakobsdottir, M., Steinberg, S., Pálsson, S., Jonasson, F., Sigurgeirsson, B., Thorisdottir, K., Ragnarsson, R., Benediktsdottir, K. R., Aben, K. K., … Stefansson, K. (2007). Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature Genetics*, *39*(12), 1443–1452.

Tedja, M. S., Wojciechowski, R., Hysi, P. G., Eriksson, N., Furlotte, N. A., Verhoeven, V. J. M., Iglesias, A. I., Meester-Smoor, M. A., Tompson, S. W., Fan, Q., Khawaja, A. P., Cheng, C.-Y., Höhn, R., Yamashiro, K., Wenocur, A., Grazal, C., Haller, T., Metspalu, A., Wedenoja, J., … Klaver, C. C. W. (2018). Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. *Nature Genetics*, *50*(6), 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*(3), 328–330.

Valverde, P., Healy, E., Sikkink, S., Haldane, F., Thody, A. J., Carothers, A., Jackson, I. J., & Rees, J. L. (1996). The Asp84Glu variant of the melanocortin 1 receptor (MC1R) is associated with melanoma. *Human Molecular Genetics*, *5*(10), 1663–1666.

Zhu, G., Duffy, D. L., Eldridge, A., Grace, M., Mayne, C., O'Gorman, L., Aitken, J. F., Neale, M. C., Hayward, N. K., Green, A. C., & 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*(2), 483–492.

Zhu, G., Evans, D. M., Duffy, D. L., Montgomery, G. W., Medland, S. E., Gillespie, N. A., Ewen, K. R., Jewell, M., Liew, Y. W., Hayward, N. K., Sturm, R. A., Trent, J. M., & 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: The Official Journal of the International Society for Twin Studies*, *7*(2), 197–210.

GILLESPIE N. A., WHITFIELD J. B., WILLIAMS B., HEATH A. C., MARTIN N. G. The relationship between stressful life events, the serotonin transporter (5-HTTLPR) genotype and major depression, Psychol Med 2005: 35: 101-111. (cites = 381)

Nick Martin

Nathan Gillespie

I first met Nick Martin almost 21 years ago when he took me on as a PhD student in his Genetic Epidemiology unit in what was then simply known as the Queensland Institute of Medical Research. Before meeting Nick, all I knew was that he was a biologist, but it became quickly apparent that his intellectual curiosity was extremely broad and encompassed many clinical and non-clinical phenotypes. Looking back over the last two decades, it is clear to me that his success in recruiting large samples of twins, his emphasis on broad but high-quality phenotyping and repeated sampling, his foresight to begin collecting genotypes and maintain genotyping efforts, his obtaining the resources for data storage and facilitating data access, his investment in his students, his ability to attract experts in genetics, his constant encouragement and his collaborative and convivial spirit have provided the launchpad for many scientific careers, including my own. Moreover, his contributions to science have given us all an invaluable resource and treasure trove of genetically and environmentally informative human data with which to tackle some of the most compelling questions regarding our nature.

Present during my first meeting with Nick was his colleague and long-time collaborator, Professor Ian Hickie. They were both keen to explore the genetic aetiology of somatic distress and its relationship to internalizing symptoms and needed a student to do the work. Under Nick's supervision, my first manuscript explored the internalizing symptoms of the DSSI and SCL scales from which we extracted the factors of somatic distress, anxiety and depression (2). Applying the Classic Mx software program to the summary polychoric correlations, we were able to show using weighted least squares that while there was significant genetic overlap between these three dimensions, there was still evidence of distinct genetic influences on somatic distress (3). The broader implication of these findings was that for individuals suffering from chronic fatigue syndrome, their symptoms of somatic distress were not entirely the same as those of depression. Several years later, Nick and I modelled the direction of causation between these same internalizing dimensions and measures of parental bonding. It was and remains a novel use of twin data, which demonstrates how under certain conditions, cross-sectional data can be effectively used to test competing causal hypotheses. The results when published (5) earned us the 2003 Fulker Award and 'very nice bottle of wine' (which Nick selected!) at the BG meeting in Aix-en-Provence. Thanks to Nick, my interest in modelling causality has persisted along with my preference for pricey Châteauneuf-du-Papes.

Not only do longitudinal data trump cross-sectional studies for testing causal hypotheses but such data provide opportunities for testing compelling developmental hypotheses. The methods for modelling genetic and environmental changes over time had been well described by Eaves et al. and others (1). However, thanks to Nick's decades-long and persistent efforts to ascertain and re-sample his twin cohorts, we were in a position to begin applying and testing developmental hypotheses to well-powered samples with repeated measures. Using data from his Brisbane Longitudinal Twin Study we were able to demonstrate the emergence of different genetic and environmental influences in adolescent and teenage personality at ages 12, 14 and 16 (6). Using the repeated measures from his two adult twin cohorts, we demonstrated that the genetic risks in non-clinical symptoms of depression and anxiety at age 20 were, by and large, enduring well into an individual's 70s (7). Incidentally, we also identified 'innovative' or additional sources of genetic variance in the 30s and 40s for anxiety and depression respectively.

Regarding the genetics of depression, it was Nick who alerted me to Caspi et al's (4) report investigating the interaction between stressful life events and the serotonin transporter genotype as a predictor of major depression. Nick realized that he had isomorphic data to validate this landmark finding using his Australian adult twins and we immediately set to work. Regardless of whether our results were based on binary logistic or ordinal regression analyses, we were the first to find no corroborating evidence supporting a main effect of 5-HTTLPR, or an interaction between the 5-HTTLPR genotype and stressful life events on major depression (8). In the decade that followed, the field entirely recalibrated its views towards candidate genes and candidate gene by environmental interactions concerning complex behaviours (9).

In 2004 I left Nick's laboratory at QIMR for what was to be a brief two-year visit to the Virginia Institute for Psychiatric and Behavior Genetics (VIPBG) in Richmond, Virginia where Nick himself had worked during the early 80s with the likes of Lindon Eaves his PhD supervisor, Kenneth Kendler, Michael Neale, John Hewitt and Andrew Heath among others. And although two years have somehow (and very quickly) turned into 16 years, I have had the great fortune of maintaining my collaboration with Nick on grant applications, numerous projects and many more publications. Perhaps the greater fortunes include the many meals I have had at the Martin household dinner table or better still, his being my best man at my wedding in 2017. I am deeply indebted to his support, confidence in me and abiding friendship. I will remain forever grateful for the opportunities that his tutelage, which has never really stopped, continues to give me.

References
1. EAVES L., EYSENCK H. J., MARTIN N. G. Genes, Culture, and Personality: An Empirical Approach London: Academic Press; 1989.
2. GILLESPIE N., KIRK K. M., HEATH A. C., MARTIN N. G., HICKIE I. Somatic distress as a distinct psychological dimension, Soc Psychiatry Psychiatr Epidemiol 1999: 34: 451-458.
3. GILLESPIE N. A., ZHU G., HEATH A. C., HICKIE I. B., MARTIN N. G. The genetic aetiology of somatic distress, Psychol Med 2000: 30: 1051-1061.
4. CASPI A., SUGDEN K., MOFFITT T. E., TAYLOR A., CRAIG I. W., HARRINGTON H. et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene, Science 2003: 301: 386-389.
5. GILLESPIE N. A., ZHU G., NEALE M. C., HEATH A. C., MARTIN N. G. Direction of causation modeling between cross-sectional measures of parenting and psychological distress in female twins, Behav Genet 2003: 33: 383-396.
6. GILLESPIE N. A., EVANS D. E., WRIGHT M. M., MARTIN N. G. Genetic simplex modeling of Eysenck's dimensions of personality in a sample of young Australian twins, Twin Res 2004: 7: 637-648.
7. GILLESPIE N. A., KIRK K. M., EVANS D. M., HEATH A. C., HICKIE I. B., MARTIN N. G. Do the genetic or environmental determinants of anxiety and depression change with age? A longitudinal study of Australian twins, Twin Res 2004: 7: 39-53.
8. GILLESPIE N. A., WHITFIELD J. B., WILLIAMS B., HEATH A. C., MARTIN N. G. The relationship between stressful life events, the serotonin transporter (5-HTTLPR) genotype and major depression, Psychol Med 2005: 35: 101-111.
9. DUNCAN L. E., KELLER M. C. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry, Am J Psychiatry 2011: 168: 1041-1049.

(Nick Martin and Nathan Gillespie 2004)

DM Evans, IH Frazer & NG Martin (1999). Genetic and environmental causes of variation in basal levels of blood cells. *Twin Res*, 2(4), 250- 257. (cites = 188)

It's in the bloody genes!

David Evans

I first heard about Nick Martin in 1998. I had come to his lab to have lunch with my former honours supervisor, Dr Margie Wright who had recently moved from the University of Queensland psychology department to work in Nick's group at the Queensland Institute of Medical Research. I remember asking her what she was working on, and seeing some complicated looking figures (path diagrams) and equations (LISREL notation) in a thick white book (Neale and Cardon). Although I had literally done no genetics at university, I remember thinking that the combination of statistics and biology looked interesting and might suit me.
A few weeks later I was back to meet Nick for the first time to discuss the possibility of doing a PhD in his lab. I was working as a medical representative for a pharmaceutical company at the time, having left university temporarily after a particularly hectic honours year, and wanted to earn some money for a while, and do something outside of academia. "God! Why the hell are you doing that?" Nick asked me in what I would soon come to recognise as classic Martin fashion, when I first met him. "Why don't you go and work on an oil rig off the coast for a year?" Apparently he also later commented to Margie about how I seemed nice enough, but wasn't a massive fan of my baseball cap (which he swore black and blue lowered the IQ of the wearer by at least ten points). Despite his reservations over my choice of fashion, Nick accepted me into his lab shortly thereafter, in what would later turn out to be a major turning point in my life- although sadly for both of us, not my wardrobe…

The PhD project I chose was a genetic study of blood cell concentrations in Nick's adolescent twin cohort. Unbeknownst to me at the time, many years previously Nick had had a meeting with Ian Frazer, the inventor of the Gardasil vaccine for human papilloma virus and cervical cancer (back then it would still be some years before the vaccine had FDA approval and Ian had attained scientific superstar status). Nick had had the incredible foresight to make sure that the blood samples taken from the twins were sent over to Ian at the Princess Alexandra Hospital (where incidentally I am now based) for a full blood count and lymphocyte subsets analysis. The result was the largest genetically informative dataset of blood cell measures in the world at that time.

The maiden paper from my PhD was the first large scale study of the heritability of blood cell counts [1]. In it we showed that the concentration of most blood cells was highly heritable, despite the considerable inter-individual and circadian variation that characterized such measures. The heritability study was published in *Twin Research and Human Genetics* and strangely enough, even though the field has moved on, is still one of the most cited papers in its area- although perhaps not for the reasons we expected. The main reason in fact is that apparently the manuscript has become a teaching aid for many aspiring students of behavior genetics who wish to use the classical twin study to conduct heritability analyses. Buried within its pages is a description of the procedure for testing the equality of means, variances and covariances across the different sexes, birth orders, and zygosities- which Nick dutifully drummed into my head during the first few months of my PhD. The paper is used as a teaching aid at the introductory Boulder Workshop and apparently at others around the world also. Indeed the paper always gets a citation bounce every alternate year because of its regular appearance (I have christened this effect on citation counts, the "Medland Effect").

In 2004, we published the first linkage studies of blood cell traits, including one of the first papers to use multivariate QTL linkage analysis in order to detect complex trait loci [2-4], and

several years later genome-wide association meta-analyses [5-8] - although by this time I had long left the lab and Manuel Ferreira and others had taken over the lead on this work. It always makes me smile that these later papers have made it into some of the most prestigious scientific journals (*Cell*, *Nature*, *American Journal of Human Genetics*), yet I remember very clearly an incident from a CRC conference where a senior Australian academic (who I will not name) got up and after a tirade of 3 minutes tried to skewer me, a fresh faced second year PhD student, about what the point was in analysing the genetics of blood cells, because we "knew everything about them already" and I was wasting my and everyone else's time.

Needless to say, Nick disagreed- and quite clearly so did *Cell*, *Nature* and many other top tier journals as it turned out.

Happy 70th birthday Nick. Thanks for your generosity and for being a supportive supervisor and mentor. I look forward to collaborating with you for many years to come.

David Evans
14/2/2020

References

1. Evans, D.M., I.H. Frazer, and N.G. Martin (1999). Genetic and environmental causes of variation in basal levels of blood cells. *Twin Res*, 2, 250-7.
2. Evans, D.M., et al. (2004). A major quantitative trait locus for CD4-CD8 ratio is located on chromosome 11. *Genes Immun*, 5, 548-52.
3. Evans, D.M., et al. (2004). Major quantitative trait locus for eosinophil count is located on chromosome 2q. *J Allergy Clin Immunol*, 114, 826-30.
4. Evans, D.M., et al. (2004). Multivariate QTL linkage analysis suggests a QTL for platelet count on chromosome 19q. *Eur J Hum Genet*, 12, 835-42.
5. Ferreira, M.A., et al. (2009). Sequence variants in three loci influence monocyte counts and erythrocyte volume. *Am J Hum Genet*, 85, 745-9.
6. Ferreira, M.A., et al. (2010). Quantitative trait loci for CD4:CD8 lymphocyte ratio are associated with risk of type 1 diabetes and HIV-1 immune control. *Am J Hum Genet*, 86, 88-92.
7. Gieger, C., et al. (2011). New gene functions in megakaryopoiesis and platelet formation. *Nature*, 480, 7376, 201-8.
8. van der Harst, P., et al. (2012). Seventy-five genetic loci influencing the human red blood cell. *Nature*, 492, 369-75.

(Brisbane Twin Study)

Medland SE et al. (2009). Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet*, 85, 750-5. (cites = 229)

Curly questions

Sarah Medland

I've had the great pleasure of working with Nick in one way or another since 2002. Over the years Nick has worked on almost everything and at one stage Lon Cardon actually worked out the multiple testing burden of *GWAS for everything* with the aim of tackling Nick with this at the conclusion of one of his *ra-ra* talks at Boulder. While much of this work has important implications for health and well-being, some has been less serious in nature.

One particularly memorable project focused on the genetics of hair curl and resulted in a GWAS study published in the *American Journal of Human Genetics* in 2009 (Medland et al. 2009). This paper was actually designed to be a baseline paper in which we could report the details of the quality control and imputation of our GWAS data for almost 5,000 individuals that had been accumulated over seven waves of SNP based genotyping. Despite this, with the usual GenEpi style, we decided that the GWAS needed to use all the information available in our data which had been collected on a three-point scale, and use data from all available individuals which meant that we needed to model relatedness. This of course meant that the GWAS had to be run in (classic) Mx. After a great deal of work, extracting the SNP data, merging it with phenotypes, running the GWAS in the context of an AE model and outputting the results was streamlined down to ~30 seconds per SNP. Thankfully this work predated the 1000 genome project, so GWAS in the three cohorts for 2.4 million SNPs only used ~60,000 CPU hours. Our major finding was an association with the 1q21 region which included a coding, nonsynonymous variant located in the third exon of *TCHH*. The variant was only seen in European populations, and was associated with straighter hair, explaining around 6% of the variance in hair shape. While we were pretty happy about this, we were under no illusions as how important this finding was in the greater scheme of things which made what happened next even more surprising.

About a month after the paper was published, Nick was emailed a couple of questions by a UK based journalist. As it turns out the key question in that email was: *"I also wondered if you could tell me what sort of application this discovery could have in future? Might it be possible to influence the curliness of hair by different means? Could there be an alternative to heated hair straighteners?"*. A follow-up email pressed on: *Could you develop new 'treatments' (like a pill?) which would make hair curlier or straighter, rather than treating the hair directly?* To which the reply that was sent "*potentially, yes*".

We were somewhat unprepared for the deluge of media attention that descended on us a few days later when the UK *Daily Telegraph* announced to the world that we had developed a pill to straighten hair. Arriving at work the next morning we dodged tv crews from three different stations and tried to explain what had happened to the QIMR media team. The local journalists were irate that we had not given the breaking story to a local media outlet. It was unfortunately a slow news week and thanks to the view and lift model of journalism, the story spread. As shown in Fig1 below, even now 10 years later a google search for *"Nick Martin" curly hair pill* yields ~10,500 hits and the interest from both media and individuals has never gone away. Over the last 10 years we've had queries from companies from *L'Oreal* to *Lush*. Ironically, this is the only GWAS finding that we did actually meet with business development with about patenting; however, the advice was, it wasn't worth it.

Fig 1   A google trends search for curly hair pill

Of course, the grass is always greener on the other side, so in European countries this magical pill was apparently able to curl hair, whilst in other countries around the world, the pill straightened hair.

Across the years my favourite query about this work came in the form of a physical letter from Brazil which arrived with six stamps on it (one of which most appropriately showed a very curly haired Costureira). The letter was written in 20pt font (mostly in caps) and included the very memorable line "I AM A DEALER IN BRAZIL", a request for exclusivity, and an intriguing question about whether the pills were for ETERNAL use.

Obviously, we missed a trick on this one Nick. Just think of what the funding line would have looked like had we actually tried to make a pill to straighten hair…

The Genetics of Reading and Language

Michelle Luciano and Timothy Bates

One of us (Bates) first met Nick Martin at the BGA meeting in Sydney. I was a student with no status in BG, but Nick's personal warmth and gregariousness welcomes all-comers. This would have its first concrete effect a dozen years later when, working in Sydney, we successfully applied for a modest NHMRC grant to study the genetics of reading, testing predictions from the leading "Dual Route Cascaded" computational model of reading. We had proposed collecting our own twin sample: As most reading this will be aware, to ascertain, zygosity test, and phenotype 500 pairs of twins was a daunting prospect (though not to Nick, who'd done just this for his - wait for it - undergraduate thesis in Adelaide!). An email became a phone call, and a trip up to Brisbane, which soon morphed into our training the professional testers who Nick had assembled on the subtleties of assessing nonword pronunciation over the telephone! Soon enough the first twin study from this project was published (Bates el al., 2004). It addressed, with key collaborators Anne Castles and Max Coltheart, aspects of the Dual Route Cascade model of reading and showed that the genetic factor structure mimicked the phonological and lexical pathways to reading aloud and not a connectionist model that was also popular in cognitive science. With a grant far too small to accomplish our goals, Nick, through his generosity, encouragement, smooth management systems, efficient and warm personal relations and the support of the large team of researchers, assistants, post-docs, and PhDs, all of whom leant a hand, made it possible not just to deliver on our goal, but to get ahead and over-deliver as a series of analyses emerged. This leads us naturally into the molecular phase of the longer-term project.

The other of us (Luciano) had been lucky enough to receive Nick's red pen marks on her thesis chapter drafts despite him not being an official PhD supervisor. Nick has forever been provocative, sharp-minded, and extremely helpful, and none more so as when he encouraged me as a recent PhD graduate from his lab to get involved in a new study on the genetics of reading and language – something that I am still active in today. With data collection now complete, the reading project began to take on a different, more QIMR-aligned direction. While cognitive scientists are interested in the models, well Nick, he was interested in the genes, and, with the human genome project beginning to pay off, this was a great time to realise that interest. Nick had been successful in attaining funding for microsatellite genotyping in the twin adolescent sample, and so onto linkage analysis we went. Now with my own funding as a research fellow on the project I performed my first (and reading ability's first) linkage analysis! Just as Nick was helpful in guiding me through twin analyses during my PhD, he continued to provide support in this new postdoc period. Nick wasn't helping with the finer details of the analysis, but he was (and still is) always on top of the latest research developments in statistical genetics (he'll point you to the right paper!). Our linkage study (Bates et al, 2007) was the first of reading ability in an unselected sample, and while we didn't expect too much in way of significant results, we contributed to quite a number of replications of candidate dyslexia genes and QTLs in our sample. One of our findings was that evidence for replication was always stronger when we removed IQ variance from our reading measure.

Nick who always looked to the future and kept pace with technological advances soon had funding for genome-wide SNP genotyping. And next, came the first GWAS of reading ability in an unselected sample, well two samples actually. Nick always encouraged collaboration and with that in mind, we got ALSPAC on board to contribute to a GWAS meta-analysis of reading, spelling and language traits (Luciano et al., 2013). Again, the sample size was relatively small,

and we didn't find much. GWAS samples of children are always, by comparison, going to be smaller than those of adults, so why not study reading and language in adults? Which is exactly what we have recently done with Nick's ongoing support, another successful phone interview study on Australian Twins that is supporting a proof-of-principle that GWAS of adult reading and language phenotypes will help us understand their disorder in development. We collected educational attainment data in this study, but could not collect IQ data (controlling for nonverbal IQ allows the reading and language measures to be much more sensitive to genetic effects – it seems that not everything is IQ, Nick). Where we've come to date in this ongoing story is not possible to compress into a single trait – not even one as "general" as IQ, which both of us worked on with Nick. Nor either, it seems to us, is it simply high ability that led Nick to ask all those years ago "What genes do you think make some kids get better grades?" and build this into a diverse scientific legacy involving hundreds of papers and collaborations across topics as diverse as attitudes, methods, Alzheimer's, baldness, twinning, and skin cancer. This Opera-loving, poetry and essay reading, adventurous, bon vivant manages to combine curiosity, boldness, warmth, interest in both big societally important questions, the openness to adopt new methods, ambition and collaborative skill to bring into being the infrastructure and samples needed for this research. It's all too rare, and we are grateful: Thank you Nick, for these, and for many more years to come!

References

Bates, T. C., Castles, A., Coltheart, M., Gillespie, N., Wright, M., & Martin, N. G. (2004). Behaviour genetic analyses of reading and spelling: a component processes approach. Australian Journal of Psychology, 56(2), 115-126.

Castles, A., Bates, T. C., Coltheart, M., Luciano, M., & Martin, N. G. (2006). Cognitive modelling and the behaviour genetics of reading. Journal of Research in Reading, 29(1), 92-103. Retrieved from <Go to ISI>://000234839800006

Bates, T. C., Castles, A., Luciano, M., Wright, M., Coltheart, M., & Martin, N. (2007). Genetic and environmental bases of reading and spelling: A unified genetic dual route model. Reading and Writing, 20(1-2), 147-171. doi:10.1007/s11145-006-9022-1

Bates, T. C., Lind, P., Luciano, M., Duffy, D., Castles, A., Wright, M. J., & Martin, N. G. (2007, 18 September, 2007). DCDC2 and DYXC1 related to normal variance in reading. Paper presented at the International Society for Psychiatric Genetics, New York.

Luciano, M., Lind, P. A., Duffy, D. L., Castles, A., Wright, M. J., Montgomery, G. W., . . . Bates, T. C. (2007). A haplotype spanning KIAA0319 and TTRAP is associated with normal variation in reading and spelling ability. Biological Psychiatry, 62(7), 811-817. doi:10.1016/j.biopsych.2007.03.007

Bates, T. C., Luciano, M., Lind, P. A., Wright, M. J., Montgomery, G. W., & Martin, N. G. (2008). Recently-derived variants of brain-size genes ASPM, MCPH1, CDK5RAP and BRCA1 not associated with general cognition, reading or language. Intelligence, 36(6), 689-693. Retrieved from <Go to ISI>://000261410000020

Lind, P., Luciano, M., Duffy, D., Castles, A., Wright, M. J., Martin, N. G., & Bates, T. C. (2009). Dyslexia and DCDC2: normal variation in reading and spelling is associated with

DCDC2 polymorphisms in an Australian population sample. European Journal of Human Genetics, 18(6), 668-673. doi:10.1038/ejhg.2009.237

Martin, N. W., Hansell, N. K., Wainwright, M. A., Shekar, S. N., Medland, S. E., Bates, T. C., . . . Wright, M. J. (2009). Genetic covariation between the Author Recognition Test and reading and verbal abilities: what can we learn from the analysis of high performance? Behav Genet, 39(4), 417-426. doi:10.1007/s10519-009-9275-y

Bates, T. C., Lind, P. A., Luciano, M., Montgomery, G. W., Martin, N. G., & Wright, M. J. (2010). Dyslexia and DYX1C1: deficits in reading and spelling associated with a missense mutation. Molecular Psychiatry, 15(12), 1190-1196. doi:10.1038/mp.2009.120

Bates, T., Luciano, M., Castles, A. *et al.* Replication of reported linkages for dyslexia and spelling and suggestive evidence for novel regions on chromosomes 4 and 17. *Eur J Hum Genet* **15,** 194–203 (2007). https://doi.org/10.1038/sj.ejhg.5201739

Luciano, M., Montgomery, G. W., Martin, N. G., Wright, M. J., & Bates, T. C. (2011). SNP sets and reading ability: testing confirmation of a 10-SNP set in a population sample. Twin Res Hum Genet, 14(3), 228-232. doi:10.1375/twin.14.3.228

Gialluisi, A., Newbury, D. F., Wilcutt, E. G., Olson, R. K., DeFries, J. C., Brandler, W. M., . . . Fisher, S. E. (2014). Genome-wide screening for DNA variants associated with reading and language traits. Genes Brain Behav, 13(7), 686-701. doi:10.1111/gbb.12158

Ritchie, S., Bates, T., & Plomin, R. (2013). Influences of reading ability on the development of intelligence: Evidence from a longitudinal monozygotic twin-discordant design. Behavior Genetics, 43(6), 538-538. Retrieved from <Go to ISI>://WOS:000326895400121

Luciano, M., Gow, A. J., Pattie, A., Bates, T. C., & Deary, I. J. (2018). The Influence of Dyslexia Candidate Genes on Reading Skill in Old Age. Behav Genet, 48(5), 351-360. doi:10.1007/s10519-018-9913-3

Nyholt DR, Low SK, Anderson CA, Painter JN, Uno S, Morris AP, Macgregor S, Gordon SD, Henders AK, Martin NG *et al.* (2012). Genome-wide association meta-analysis identifies new endometriosis risk loci. Nat Genet 44, 1355-9. (cites = 211)

The Genetics of Endometriosis

G.W. Montgomery

The genesis of the endometriosis mapping project lay in a survey on women's health in twins conducted by Susan Treloar and Nick Martin. One rather surprising result from the analyses of these data was the relatively high heritability for hysterectomy. Of course, the reason for this is the relatively high heritability for endometriosis and uterine fibroids, two of the main risk factors for hysterectomy. Sue and Nick were considering follow up studies on the genetics of these two main risk factors.

Not for the first time, a chance dinner party tipped the balance on which project to start with. A collaboration with local Gynaecologist Dr Dan O'Connor was started, and with his support, the project on genetic risk factors for endometriosis began. The project was helped in the initial stages by a generous donation from a family with a history of the disease. Not long afterwards a successful application to the Commonwealth Government provided funding for the Cooperative Research Centre for the Discovery of Common Human Disease and the endometriosis project became a flagship project for the Disease CRC. This major injection of funding enabled recruitment of a large cohort of women with surgically confirmed disease and the genotyping of genetic markers for linkage mapping. The sample and dataset recruited by Sue and Nick remains one of the largest samples in the world with surgically confirmed disease and a cornerstone of the continuing efforts to map genetic risk factors for endometriosis.

The project had started before I arrived in Brisbane. At the time I was working in the Biochemistry Department at the University of Otago in New Zealand, and running an animal gene mapping program. We had spent the previous 10 years mapping and cloning genes for dizygotic twinning. The fact this project could now be done in a matter of months using next generation sequencing is one measure of progress. I had been corresponding with Nick about the twinning projects and visited Brisbane for the first time on my way home from a conference in Darwin. Nick's program was expanding and I was looking for new challenges so I packed up and moved to Brisbane to join Nick's group.

Sue and her team were working hard identifying women with endometriosis, obtaining individual surgical reports to confirm diagnosis and arranging for collection of blood samples. I took over the laboratory component, receiving and processing the blood samples, extracting DNA, and preparing sets of samples to send away for genotyping. Over the course of the project we collected blood samples from over 9000 women with endometriosis and their families. This coincided with several large NIH grants to collect blood and DNA samples in other projects including the twins. I had run some big studies in my animal genomics program, but not on this scale or with the same level of detail. It was a hectic few years developing systems while handling the increasing volume of incoming samples as recruitment ramped up in multiple studies. At the peak, we were recruiting and processing samples from 8000 participants each year.

There were some funny incidents along the way like the day early on, when we saw a delegation of senior Institute staff going into Nick's office. It appeared they were not aware that Nick had a laboratory. Ordering laboratory supplies had triggered a delegation to ensure we had regulatory approvals (we had) and were paying our share of laboratory costs (perhaps not so much). They were also the days of the long roles of brown paper used to map out the needs, database developments and timelines for the laboratory so we could meet the increasing demands. In the end, the formidable team Nick assembled for both recruitment and laboratory processing was highly successful. The scale and quality of data and sample processing achieved were second to none and provided the platform for the many successes to follow in projects including our studies in endometriosis.

Following the major recruitment drive in the endometriosis project, samples were prepared and sent to the Australian Genome Research Facility (AGRF) in Melbourne for microsatellite genotyping funded by the CRC. This was the largest genotyping project undertaken by the AGRF. At the time of course, it

was emerging that early hope for linkage studies was optimistic and larger samples were likely to be needed for success. In parallel with our studies, the group in Oxford led by Stephen Kennedy was also conducting linkage studies. We began discussing collaboration though the CRC and the respective commercial partners for both Australia and Oxford. The terms of collaboration were agreed and the deal was signed in Singapore, perhaps in the bar of Raffles Hotel, but that could be apocryphal.

The next step was the great unveiling of results by the two sides. That took place at a meeting in QIMR when a delegation from Oxford flew out to Australia. It was a memorable day because it was the day the planes flew into the Twin Towers in New York. I was woken early by a phone call from family in New Zealand telling me to turn on the television and like so many others watched events unfold in real time. By the time we arrived at QIMR for the great unveiling, we were all shocked, not least because our colleagues were due to fly out from Australia in a few days and the immediate future of air travel looked very uncertain.

Nevertheless, we addressed the business at hand and the two sides presented fascinating results with both groups showing evidence of linkage on chromosome 9 near CDKN2A, a region we were very familiar with from our melanoma studies. The genotyping was still to be completed by each group and when the final results were analysed, the evidence for linkage to this region had faded away. There was at best marginal evidence for linkage on a region of chromosome 10. These results were published in the American Journal of Human Genetics in 2005 (Treloar *et al.*, 2005). We obtained an NIH grant to conduct follow up genotyping across this region. However, a curious fact is that the fine mapping and subsequent GWAS results have not provided evidence of association on chromosome 10 (Painter *et al.*, 2011b), but there is association in the region of the original linkage evidence on chromosome 9 (Nyholt *et al.*, 2012).

New approaches were needed and this corresponded with the development of high-throughput genotyping chips and GWAS. We were successful with NHMRC and Wellcome Trust Funding for SNP genotyping of samples from Brisbane and Oxford. At QIMR we conducted replication genotyping in a sample from the Nurses' Health Study from Boston. We published the first GWAS study for European women in 2011 (Painter *et al.*, 2011a). We identified one novel region on chromosome 7 and replicated a result published by a Japanese group the year before of association on chromosome 1 (Painter *et al.*, 2011a). We contacted the Japanese group and the next year completed a meta-analysis of data from the two groups, replicating both earlier results and novel associations for a further 5 genomic regions (Nyholt *et al.*, 2012). The studies continued with the International Endometriosis Genetics Consortium that has greatly expanded from the early days with just Brisbane and Oxford. In 2017 we published a meta-analysis reporting a total of 14 "hits" and this has expanded to 44 "hits" with analysis of 60,000 cases likely to be published this year. The research continues to be expanded in other ways. This is but one example of how the foundations were firmly laid by Nick's drive and enthusiasm for the project.

References
Nyholt DR, Low SK, Anderson CA, Painter JN, Uno S, Morris AP, Macgregor S, Gordon SD, Henders AK, Martin NG *et al.* (2012) Genome-wide association meta-analysis identifies new endometriosis risk loci. Nat Genet 44, 1355-9.

Painter JN, Anderson CA, Nyholt DR, Macgregor S, Lin J, Lee SH, Lambert A, Zhao ZZ, Roseman F, Guo Q *et al.* (2011a) Genome-wide association study identifies a locus at 7p15.2 associated with endometriosis. Nat Genet 43, 51-4.

Painter JN, Nyholt DR, Morris A, Zhao ZZ, Henders AK, Lambert A, Wallace L, Martin NG, Kennedy SH, Treloar SA *et al.* (2011b) High-density fine-mapping of a chromosome 10q26 linkage peak suggests association between endometriosis and variants close to CYP2C19. Fertil Steril 95, 2236-40.

Treloar SA, Wicks J, Nyholt DR, Montgomery GW, Bahlo M, Smith V, Dawson G, Mackay IJ, Weeks DE, Bennett ST *et al.* (2005) Genomewide linkage study in 1,176 affected sister pair families identifies a significant susceptibility locus for endometriosis on chromosome 10q26. Am J Hum Genet 77, 365-76.

Migraine, Human Genetics and a Passion for Science.

Dale R Nyholt

My first interactions with Professor Nicholas ("Nick") G Martin occurred at the end of my PhD during the inaugural Australasian Human Gene Mapping ('GeneMappers') Meeting held in Thredbo (NSW, Australia) in early February 1999—a couple of weeks before I flew to New York to begin my first postdoc in Jurg Ott's Statistical Genetics Laboratory at Rockefeller University.

I vividly recall (okay, foggily recall [there may have been alcohol involved]) lively discussions on gene mapping and encouragement to contact him should I want to continue my research career upon returning to Australia. Less than two years later and I was sitting in Nick's office with a lovely view of the Brisbane skyline finalising my NHMRC early career (Peter Doherty) fellowship application. From these early interactions, I learnt that Nick always spoke his mind (often with wild abandon), but he was always motivated by a desire to perform good science. It is this infectious passion for science that attracts and inspires those around him.

My fellowship application was successful and a few months later I began my journey as Nick's colleague and collaborator within his Genetic Epidemiology Laboratory.

Nick created and continues to maintain a world-class research environment that is rich in data, expertise and excellence. I will always be grateful for the opportunity to learn and benefit from this environment. Indeed, although I was initially attracted to Nick's lab to ask new and deep questions that extended my PhD research on migraine genetics, I was able to both lead and contribute to hundreds of genetic studies comprising dozens of traits- prominent examples include depression (Yang et al., 2018), endometriosis (Nyholt et al., 2009; Sapkota et al., 2017), leukocyte telomere length (Broer et al., 2013), male pattern baldness (Nyholt, Gillespie, Heath, & Martin, 2003), obesity (Locke et al., 2015; Rahmioglu et al., 2015), and twinning (Mbarek et al., 2016).

Our collaborative research has produced important advances and paradigm changes. For example, one of our first publications rebuked the widely accepted opinion that common baldness was an autosomal dominant phenotype in men and an autosomal recessive phenotype in women. In this first large-scale study of 476 monozygotic (MZ) and 408 dizygotic (DZ) male twin pairs aged we estimated a heritability of 0.81 (95% CI: 0.77–0.85) and indicated that additive genetic effects play a major part in the progression of common male hair loss (Nyholt et al., 2003).

Similarly, our migraine research applied latent class and twin genetic analyses to identify subgroups of migraine sufferers and show the existence of a severity continuum, where migraine with aura is more severe, but not, as previously thought, aetiologically distinct from migraine without aura (Nyholt et al., 2004). This research attracted international attention and led to high impact migraine collaborations that persist today. Indeed, this research, together with Nick's extensive network of international twin/genetic researchers led to the co-founding of the International Headache Genetics Consortium (IHGC), which brought together headache geneticists and clinicians from around the globe, to conduct numerous large-scale genetic studies on migraine (Anttila et al., 2006; Anttila et al., 2008; Ligthart, Boomsma, Martin, Stubbe, & Nyholt, 2006; Ligthart et al., 2008; Mulder et al., 2003; Nyholt et al., 2005; van den Maagdenberg, Nyholt, & Anttila, 2019).

With advancing genotyping technology, our migraine research was at the forefront of genetic association studies. Our 2008 IHGC publication showed that contrary to the leading hypothesis at the time, ion transport genes—implicated in Familial Hemiplegic Migraine (FHM), a Mendelian subtype of migraine with aura (MA) associated with hemiparesis—did not play a major role in the common forms of migraine (Nyholt et al., 2008). Our research also showed that despite the female:male prevalence ratio of >2:1, female and male migraineurs are not genetically distinct (Mulder et al., 2003; Nyholt et al., 2004; Nyholt et al., 2015). These advances were crucial to the design and execution of subsequent well-powered genetic studies of migraine—all led by the IHGC.

In 2016, we published the largest ever genetic study of migraine (involving 59,674 migraine cases and 316,078 controls) and identified 44 (34 new) risk variants for migraine (Gormley et al., 2016). Most prominently, this research provided valuable insight into migraine pathophysiology, by indicating vascular dysfunction to be a primary mechanism underlying migraine. This is important, because there is a long running debate about whether migraine is a disease of vascular dysfunction, or a result of neuronal dysfunction with secondary vascular changes. This paper's Altmetric attention score is in the top 0.02% of all research outputs ever tracked. Moreover, the results from this study allows polygenic risk score (PRS) analyses in migraine risk prediction to identify/quantify comorbidities, endophenotypes, and drug responses; and paves the way to develop relevant vascular cellular models of migraine that are required to understand the molecular mechanisms of migraine and develop new drugs. Indeed our migraine PRS was able to identify subgroups of patients likely to respond to triptans (an acute migraine drug), providing the first step toward precision medicine in migraine (Kogelman et al., 2019).

The above highlights are but a few of the many that I've been fortunate to share with Nick over the past 20 years, and I hope to share many more. As we celebrate Nick's 70th Birthday (besides from mentally noting his fitting Platinum Jubilee themed hair colour), I fondly reflect on the countless discussions, opportunities and accomplishments we have shared, and I marvel at the amazing legacy he continues to build for current and future generations. Cheers Nick, you are truly a unique and special individual.

Yours Sincerely,

Dale.

References

Anttila, V., Kallela, M., Oswell, G., Kaunisto, M. A., Nyholt, D. R., Hamalainen, E., . . . Palotie, A. (2006). Trait components provide tools to dissect the genetic susceptibility of migraine. *American journal of human genetics, 79*(1), 85-99. doi: 10.1086/504814

Anttila, V., Nyholt, D. R., Kallela, M., Artto, V., Vepsalainen, S., Jakkula, E., . . . Palotie, A. (2008). Consistently replicating locus linked to migraine on 10q22-q23. *American journal of human genetics, 82*(5), 1051-1063. doi: 10.1016/j.ajhg.2008.03.003

Broer, L., Codd, V., Nyholt, D. R., Deelen, J., Mangino, M., Willemsen, G., . . . Boomsma, D. I. (2013). Meta-analysis of telomere length in 19,713 subjects reveals high heritability, stronger maternal inheritance and a paternal age effect. *European journal of human genetics : EJHG, 21*(10), 1163-1168. doi: 10.1038/ejhg.2012.303

Gormley, P., Anttila, V., Winsvold, B. S., Palta, P., Esko, T., Pers, T. H., . . . Palotie, A. (2016). Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nature genetics, 48*(8), 856-866. doi: 10.1038/ng.3598

Kogelman, L. J. A., Esserlind, A.-L., Christensen, A. F., Awasthi, S., Ripke, S., Ingason, A., . . . The International Headache Genetics Consortium. (2019). Migraine polygenic risk score associates with efficacy of migraine-specific drugs. *Neurology Genetics, 5*(6), e364. doi: 10.1212/nxg.0000000000000364

Ligthart, L., Boomsma, D. I., Martin, N. G., Stubbe, J. H., & Nyholt, D. R. (2006). Migraine with aura and migraine without aura are not distinct entities: further evidence from a large Dutch population study. *Twin research and human genetics : the official journal of the International Society for Twin Studies, 9*(1), 54-63. doi: 10.1375/183242706776403019

Ligthart, L., Nyholt, D. R., Hottenga, J. J., Distel, M. A., Willemsen, G., & Boomsma, D. I. (2008). A genome-wide linkage scan provides evidence for both new and previously reported loci influencing common migraine. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics, 147B*(7), 1186-1195. doi: 10.1002/ajmg.b.30749

Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., . . . Speliotes, E. K. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature, 518*(7538), 197-206. doi: 10.1038/nature14177

Mbarek, H., Steinberg, S., Nyholt, D. R., Gordon, S. D., Miller, M. B., McRae, A. F., . . . Boomsma, D. I. (2016). Identification of Common Genetic Variants Influencing Spontaneous Dizygotic Twinning and Female Fertility. *American journal of human genetics, 98*(5), 898-908. doi: 10.1016/j.ajhg.2016.03.008

Mulder, E. J., Van Baal, C., Gaist, D., Kallela, M., Kaprio, J., Svensson, D. A., . . . Palotie, A. (2003). Genetic and environmental influences on migraine: a twin study across six countries. *Twin research : the official journal of the International Society for Twin Studies, 6*(5), 422-431. doi: 10.1375/136905203770326420

Nyholt, D. R., Gillespie, N. A., Heath, A. C., & Martin, N. G. (2003). Genetic basis of male pattern baldness. *The Journal of investigative dermatology, 121*(6), 1561-1564. doi: 10.1111/j.1523-1747.2003.12615.x

Nyholt, D. R., Gillespie, N. G., Heath, A. C., Merikangas, K. R., Duffy, D. L., & Martin, N. G. (2004). Latent class and genetic analysis does not support migraine with aura and migraine without aura as separate entities. *Genetic epidemiology, 26*(3), 231-244. doi: 10.1002/gepi.10311

Nyholt, D. R., Gillespie, N. G., Merikangas, K. R., Treloar, S. A., Martin, N. G., & Montgomery, G. W. (2009). Common genetic influences underlie comorbidity of migraine and endometriosis. *Genetic epidemiology, 33*(2), 105-113. doi: 10.1002/gepi.20361

Nyholt, D. R., International Headache Genetics Consortium, Anttila, V., Winsvold, B. S., Kurth, T., Stefansson, H., . . . Palotie, A. (2015). Concordance of genetic risk across migraine subgroups: Impact on current and future genetic association studies. *Cephalalgia : an international journal of headache, 35*(6), 489-499. doi: 10.1177/0333102414547784

Nyholt, D. R., LaForge, K. S., Kallela, M., Alakurtti, K., Anttila, V., Farkkila, M., . . . Palotie, A. (2008). A high-density association screen of 155 ion transport genes for involvement with common migraine. *Human molecular genetics, 17*(21), 3318-3331. doi: 10.1093/hmg/ddn227

Nyholt, D. R., Morley, K. I., Ferreira, M. A., Medland, S. E., Boomsma, D. I., Heath, A. C., . . . Martin, N. G. (2005). Genomewide significant linkage to migrainous headache on

chromosome 5q21. *American journal of human genetics, 77*(3), 500-512. doi: 10.1086/444510

Rahmioglu, N., Macgregor, S., Drong, A. W., Hedman, A. K., Harris, H. R., Randall, J. C., . . . Zondervan, K. T. (2015). Genome-wide enrichment analysis between endometriosis and obesity-related traits reveals novel susceptibility loci. *Human molecular genetics, 24*(4), 1185-1199. doi: 10.1093/hmg/ddu516

Sapkota, Y., Steinthorsdottir, V., Morris, A. P., Fassbender, A., Rahmioglu, N., De Vivo, I., . . . Nyholt, D. R. (2017). Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. *Nat Commun, 8*, 15539. doi: 10.1038/ncomms15539

van den Maagdenberg, A. M. J. M., Nyholt, D. R., & Anttila, V. (2019). Novel hypotheses emerging from GWAS in migraine? *J Headache Pain, 20*(1), 5. doi: 10.1186/s10194-018-0956-x

Yang, Y., Zhao, H., Boomsma, D. I., Ligthart, L., Belin, A. C., Smith, G. D., . . . Nyholt, D. R. (2018). Molecular genetic overlap between migraine and major depressive disorder. *European journal of human genetics : EJHG, 26*(8), 1202-1216. doi: 10.1038/s41431-018-0150-2

(Using both fingers?)

Visscher PM, Medland SE, Ferreira MA, Morley KI, Zhu G, Cornes BK, Montgomery GW, Martin NG (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet*, 2(3):e41. (Cites = 492)

Musings on Visscher et al. (2006)

Peter M. Visscher

In 2004 I attended the (fourth) Australian "Genemappers" meeting in Perth. I had spent a 3-month sabbatical at the Queensland Institute of Medical Research (QIMR) the previous year, and Naomi (Wray) and I, who were working in Edinburgh at the time, were seriously considering migrating to Australia at some point – hence the interest in attending this relatively small meeting at the other side of the world. I'm pretty sure that Nick very generously paid for my travel expenses from Edinburgh to Perth (Australia, not Scotland) so that I could attend the meeting (although there might have been an element of self-interest too!).

At the meeting Nick gave a short presentation of work that Gu Zhu and he had been doing using results from genetic linkage analyses, using data from microsatellite markers on sibling pairs (mostly DZ twin pairs). Linkage analyses, i.e. the analysis of association between identity-by-descent (IBD) status at genomic loci and complex traits within families, were still population in those days – they were to be replaced by GWAS very soon afterwards. Interestingly, Gu and Nick were using the IBD estimates for a purpose that differed from the standard locus-by-locus genome scan.

The title of Nick's presentation was "Biometrical Genetics – with real data!". They had used the locus-by-locus estimates of IBD to obtain a genome-wide IBD estimate of 'realised relatedness' or 'actual relatedness' in about 900 sibling pairs, and also genome-wide coefficients of dominance. The estimate of genome-wide relationship was obtained by averaging IBD estimates across many (about 3500) points in the genome. The estimate of the mean and SD of additive and dominance relatedness were (0.5, 0.04) and (0.25, 0.04), respectively. The SD are the most interesting parameters in this context, and turn out to be spot-on with what is expected under (previously published) theory. Nick also showed results from using trait data on height (and other traits), from fitting and comparing various statistical models using genome-wide and chromosome-wide estimates of realised relationships. (Twin researchers like to perform model testing and model selection, rather than just focus on the estimation of variance components and their standard errors. I have never quite understood this, because the inference from model testing depends on the sample size and can lead to winner's curse. In addition, why would I want to calculate a p-value for narrow-sense heritability when we know that all traits that vary in the population will have some genetic variation?). There was a lack of power of the trait-based analyses, but the idea to combine realised relationships with trait data intrigued me very much.

As an aside, Nick has consistently claimed that the idea to estimate realised relationships from marker data and then perform statistical analyses for complex traits came to him while traveling on a bus in The Provence. This must be a true story, because Nick hardly ever uses public transport, let alone a bus. Nick has had other famous Road to Damascus moments in his life, not least his 180 degree turn from socialism to conservatism in his early twenties. But let's keep to the scientific eureka moments. The story of "The Great Provence Insight" was repeated many times after the events, perhaps most infamously when Nick and I (and others) were being interviewed for a major grant proposal in Australia a few years later. The interviewees had absolutely no idea what Nick was talking about scientifically (they were neither geneticists nor quantitative), but may have been envious about his (working) holiday in The Provence. In the end, we didn't get the grant, but that was most likely because of other issues.

Although Nick wasn't, to my knowledge, aware of it, theory and empirical applications of the variation in realised relatedness about the expected value (e.g. variation around 0.5 for DZ twins) goes back to the 1970s. In the 1990s, several authors had started to quantify how much of this variation could be captured with genetic markers, for example in line crosses (I worked on this in mid-late 1990s) but also in outbred populations. For complex trait data, multiple authors prior to 2004 had suggested to use estimates of relatedness from marker data and subsequently estimate genetic parameters using those estimates. However, those applications were generally in cases where the pedigree is not known, for example in ecology and evolutionary studies. Therefore, the combination of IBD-based estimation of

relatedness and complex trait analyses was novel and opened the door to address a number of interesting scientific questions using a new experimental design.

After joining QIMR in 2005, the first question I was interested in addressing using realised relatedness was the estimation of within-family additive genetic variance using sibling pairs. In a random mating population, 50% of variance is between and 50% is within families. Within-family variance is sometimes called segregation variance or the variance of Mendelian sampling terms (it is because of this variance that children have a path coefficient of $\sqrt{0.5}$ with themselves for the additive term A in an extended twin design). The association between the departures of realised relationship from its expected value (of 0.5) and trait similarity for sibling pairs can be used to estimate within-family additive genetic variance, and therefore heritability. The beauty of this experimental design is that it is free from confounding due to environmental factors and G-E correlations: we are simply comparing how similar sibling pairs are that happen to share, say, 55% of their genome IBD versus those that share, say, 45% of their genome IBD. Estimating variance in this way is the same as performing a linkage analysis with the entire genome (instead of with a single locus). Therefore, it is in theory an extremely nice design to estimate and partition genetic variation. Indeed, Nick and Gu attempted such analyses with height in 2004.

For the 2006 PLOS Genetics paper we used a sample size of 4400 pairs with marker data and 3800 pairs with both marker and data on height – a combination of DZ twins and non-twin siblings. We are now used to huge sample sizes in GWAS, but in 2005 this family-based sample size was probably the largest of its kind in the world. Indeed, it was the availability of data like those that was part of the attraction of moving to Brisbane. I had done the theory of the power of the design and realised that, unfortunately, much larger samples are needed to estimate the variance components accurately – the sampling variance of the estimate of heritability is inversely proportion to the product of sample size (N pairs) and the variance of relatedness (~0.038^2), so the standard error is proportional to $1/(0.038*sqrt(N)) \sim 26/sqrt(N)$. Our point estimate for height from segregation variance was 0.8, but with a large confidence interval ranging from 0.4 to 0.9.

Despite the large sampling variance, we believed that the paper was a nice proof of concept of a neat experimental design, and had great hopes of getting it published in a good journal. We thought that AJHG was the right journal for it but the Editor (after consultation with the Editorial Board) didn't want to send it out for review. We appealed, twice, but received a rejection every time. The only feedback we received was along the lines of "we already have a twin/pedigree design to estimate heritability, why do we need another one, in particular if it is not very powerful?". In other words, they just didn't get the novelty. We ended up in PLOS Genetics, where the referees were quantitative geneticists not working in human genetics, and the paper sailed through. Interestingly, follow-up papers (Visscher et al. 2007 and Hemani et al. 2013) did get published in AJHG, and got a fairly easy time from the referees. A reminder of the stochasticity of the system!

The subsequent papers used the same design to partition genetic variation by chromosome (2007) and included BMI as a trait (2013). The latter paper (Hemani et al. 2013) was on a total of 20,000 sibling pairs and showed clear evidence for 'genomic inflation' from linkage analysis, which is proof (as if we needed it) of the polygenicity of traits like height and BMI.

Recently, the within-family experimental design was extended for complex pedigrees by Alexander Young and Augie Kong (Young et al. 2018, Nature Genetics), who applied their method to data from deCODE. They called their method "Relatedness Disequilibrium Regression", which is a complicated but succinct way of saying that the method estimates the variance of Mendelian segregation effects. There is a renewed interest in estimating variance components using these kinds of designs because it allow the break-up of genotype-environment correlations, which are expected for traits like IQ and educational attainment. Hence, direct additive genetic effects can be estimated from within-family segregation, and these effects can be separated from parental (maternal and paternal) effects.

As with all genetic analyses, there are caveats with the estimation of genetic variance from within-family segregation. Importantly, segregation variance is not affected by non-random mating, whereas between-family variance is. Therefore, for traits like height, IQ and educational attainment, for which there is strong empirical evidence of assortative mating, the estimate of additive genetic variance from within-family estimation is expected to be lower than that inferred from the correlation between relatives, irrespective of parental ('nurturing') effects. Therefore, for traits undergoing assortative mating, the comparison of estimates of additive genetic variance (or heritability) from within and between-family experimental designs can lead to incorrect conclusions.

What's next? Although it seemed inconceivable back in 2006, it is now possible to use the within-family design on sample sizes approaching 100,000 sibling pairs, and estimate and partition genetic variance for behavioural and other complex traits with good accuracy and free from confounding factors. Those 100,000 pairs will have genome-wide SNP data from GWAS or WGS, so in principle joint within and between family analyses could be performed. Quantitative Genetics – with real data!

Genetics of Depression: sample size, sample size, sample size

Enda M Byrne, Anjali K Henders, Ian B Hickie, Christel M Middeldorp, Naomi R Wray

(Authors are listed alphabetically)

Major depressive disorder is common, affecting 10% of men and 20% of women in their lifetime. Its etiology is heterogeneous with both genetic and non-genetic risk factors. With this level of complexity, most studies of the genetics of depression call for collection of larger sample size. Nick Martin was early to recognise this, and more to the point, do something about it. As early as 1984 he published on 3,810 twin pairs[1], when prior to this, the largest published sample size for these traits was 587 twin pairs[1]. This sample size was a massive feat in the pre-digital era. Nick implemented standardised interviews (the famous 1981 white, 1989 green, 1991 yellow booklets and his success might be attributed to his attention to detail and the personal touch -*hand*-written birthday cards, prints of flowers by his mother Beryl Martin, an acclaimed water-colour artist, and always posted with a proper stamp not a postmark! In designing these questionnaires, he recognised the value of recording quantitative measures of depression related traits, such as anxiety and depression symptoms and neuroticism. Quick to adopt study designs that give best bang for buck, one study for depression and anxiety used a clinical phone interview of 2470 twins selected for their extreme scores for neuroticism in order to increase statistical power for a linkage study[2]. Given the need for an even larger sample, these data were combined with similar measures obtained in Dutch twins. Nick generously provided me (CMM) with the opportunity to come to Brisbane and analyse those data[3-5]. By the time I (NRW) joined the QIMR group in 2005 there were 12,772 twin pairs from 5,000 families with up to four longitudinal measures of neuroticism[6-8].

These data provided many important research contributions beyond the traditional ACE modelling: 1) genetic contribution to variation between people in neuroticism and depression symptoms was far more important than the shared environmental factors[3, 8], 2) despite differences between the sexes in prevalence of depression, the genetic factors were mostly shared[5], 3) that the association between childhood sexual abuse and psychopathology arises at least in part through the influence of shared familial factors on both risk of victimization and risk of psychopathology[9] 4) the relationship with post-partum lifetime depression[10]. Nick was never one to steer away from difficult or thorny problems, such as the complex relationship between marital problems and depressive symptoms [15], nicely put in this way: "The study of marital relationships and depression is not unlike a game of cat's cradle: an interactive two-person game that can produce multiple outcomes, many tied up in a frustrating knot. However, behavior genetic studies disentangle one substantial knot—the realistic possibility that genetic and environmental selection account for part of the association between marital problems and depressive symptoms…. This is because twin analyses control for measured and unmeasured genetic selection into having an unhappy marriage or feeling depressed." They showed that poor marital support is associated with depressive symptoms after accounting for the genetic factors that contribute to the cat's cradle. Nick's foresight in collection of endophenotypes and subtypes of depression such as postpartum depression[11], seasonal affective disorder[12] and insomnia[13, 14] has proven fertile ground for me (EMB) to dissect the heterogeneity of depression.

Not surprisingly, these bold and evidence-based, well-powered studies earned Nick a well-deserved international reputation and a high citation index. It was a realisation that the most highly cited researcher in psychiatry was a geneticist (and very generous and inquisitive colleague) that led me (IBH) to self-introduce and establish a now 20-year collaboration through the Brisbane Adolescent and Twin Study[16]. Adolescents aged 12-14 years were recruited over the period 1992-2016 (N~3,800 with personality data), with up to 5 waves of data collection[17], with our report on the 25Up (25 years and older) study just published[18].

Nick has always been ahead of the times, first in data collection in twin studies and then in establishing a wet lab in collaboration with Grant Montgomery for generating the genotype data for linkage studies[4, 7] and candidate gene studies[19, 20]. Of course, with the benefit of hindsight we now understand why these

studies failed (the traits are highly polygenic), but were still an important stepping stone to where we are today. Next, came the genome-wide association studies (GWAS) and Nick's QIMR samples contributed to one of the first consortium studies, the MDD2000+ study so-named because of the goal to achieve a sample of 2,000 cases[21], still massive in 2010. Our (NRW and EB) careers were boosted significantly by our entry card into international consortia provided by the QIMR depression samples. In ten short years from the MDD2000+, the international psychiatric genomics consortium has accumulated genomic data on > 175K depression cases[22]. Recognising the need for large single cohort data sets recorded not only for case-control status but with measures of a wide range of symptom, lifestyle, comorbid disorder and drug response data, Nick applied for and was awarded one of the largest NHMRC Project grants to date, AU$2.5 for the Australian Genetics of Depression Study (AGDS) (NRW, IBH and EMB are all co-Is). Nick used the skills well-learned in recruitment of twin cohorts to generate a new approach of direct-to-consumer case cohort collection, with the strong belief that individuals are capable of self-reporting and indeed can report over a longer period of time than can be achieved in clinical cohorts. After small pilot trials (don't run before you can walk), over 15,000 people completed the online surveys and provided a DNA sample in a 6-month campaign heavily using radio and TV interviews and social media (yes NGM is a very presentable media tart). The resulting data are rich and the first publications[23] are starting to come out. The UK GLAD (genetic links to anxiety and depression) study was modelled on AGDS and recruited 40,000 cases of anxiety/depression[24], providing useful reciprocal replication data.

Nick is well-known for the welcome provided to new recruits and visitors, both scientifically and socially. It is because of him, that many working in the field and quantitative and psychiatric genetics are proud to call Brisbane, Australia home (NRW, EB and CM all moved countries to work here). In the month before his 70th birthday, Nick Martin started his NHMRC Leadership 3 Fellowship, and he is fired up for 5 more years of data collection and new research results. Over his career Nick has had an uncanny talent for collecting world-recognised data sets that seem to have grown exponentially over time and are able to answer increasingly complex problems. In recognising the critical importance of sample size, particularly when it comes to genetic studies of depression, we wait with anticipation what this new funding will bring.

References

1.   Jardine R, N.G. M, Henderson AS. Genetic covariation between neuroticism and symptoms of anxiety and depression. Genetic epidemiology 1984;1:89-107.

2.   Kirk KM, Birley AJ, Statham DJ, et al. Anxiety and depression in twin and sib pairs extremely discordant and concordant for neuroticism: prodromus to a linkage study. Twin Res 2000;3:299-309.

3.   Middeldorp CM, Birley AJ, Cath DC, et al. Familial clustering of major depression and anxiety disorders in Australian and Dutch twins and siblings. Twin Res Hum Genet 2005;8:609-615.

4.   Middeldorp CM, Sullivan PF, Wray NR, et al. Suggestive linkage on chromosome 2, 8, and 17 for lifetime major depression. Am J Med Genet B Neuropsychiatr Genet 2009;150B:352-358.

5.   Middeldorp CM, Wray NR, Andrews G, Martin NG, Boomsma DI. Sex differences in symptoms of depression in unrelated individuals and opposite-sex twin and sibling pairs. Twin Res Hum Genet 2006;9:632-636.

6.   Wray NR, Birley AJ, Sullivan PF, Visscher PM, Martin NG. Genetic and phenotypic stability of measures of neuroticism over 22 years. Twin Res Hum Genet 2007;10:695-702.

7.   Wray NR, Middeldorp CM, Birley AJ, et al. Genome-wide linkage analysis of multiple measures of neuroticism of 2 large cohorts from Australia and the Netherlands. Arch Gen Psychiatry 2008;65:649-658.

8.   Lake RI, Eaves LJ, Maes HH, Heath AC, Martin NG. Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. Behav Genet 2000;30:223-233.

9.	Dinwiddie S, Heath AC, Dunne MP, et al. Early sexual abuse and lifetime psychopathology: a co-twin-control study. Psychol Med 2000;30:41-52.

10.	Treloar SA, Martin NG, Bucholz KK, Madden PA, Heath AC. Genetic influences on post-natal depressive symptoms: findings from an Australian twin sample. Psychol Med 1999;29:645-654.

11.	Byrne EM, Carrillo-Roa T, Penninx BW, et al. Applying polygenic risk scores to postpartum depression. Arch Womens Ment Health 2014;17:519-528.

12.	Byrne EM, Psychiatric Genetics Consortium Major Depressive Disorder Working G, Raheja UK, et al. Seasonality shows evidence for polygenic architecture and genetic correlation with schizophrenia and bipolar disorder. J Clin Psychiatry 2015;76:128-134.

13.	Byrne EM, Gehrman PR, Medland SE, et al. A genome-wide association study of sleep habits and insomnia. Am J Med Genet B Neuropsychiatr Genet 2013;162B:439-451.

14.	Byrne EM, Heath AC, Madden PA, et al. Testing the role of circadian genes in conferring risk for psychiatric disorders. Am J Med Genet B Neuropsychiatr Genet 2014;165B:254-260.

15.	Beam CR, Horn EE, Hunt SK, Emery RE, Turkheimer E, Martin N. Revisiting the effect of marital support on depressive symptoms in mothers and fathers: a genetically informed study. J Fam Psychol 2011;25:336-344.

16.	Wright MJ, Martin NG. Brisbane adolescent twin study: Outline of study methods and research projects. Aust J Psychol 2004;56:65-78.

17.	Couvy-Duchesne B, O'Callaghan V, Parker R, et al. Nineteen and Up study (19Up): understanding pathways to mental health disorders in young Australian twins. BMJ Open 2018;8:e018959.

18.	Mitchell BL, Campos AI, Renteria ME, et al. Twenty-Five and Up (25Up) Study: A New Wave of the Brisbane Longitudinal Twin Study. Twin Res Hum Genet 2019;22:154-163.

19.	Wray NR, James MR, Dumenil T, et al. Association study of candidate variants of COMT with neuroticism, anxiety and depression. Am J Med Genet B Neuropsychiatr Genet 2008;147B:1314-1318.

20.	Wray NR, James MR, Mah SP, et al. Anxiety and comorbid measures associated with PLXNA2. Arch Gen Psychiatry 2007;64:318-326.

21.	Wray NR, Pergadia ML, Blackwood DH, et al. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. Mol Psychiatry 2012;17:36-48.

22.	Howard DM, Adams MJ, Clarke TK, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci 2019;22:343-352.

23.	Byrne EM, Kirk KM, Medland SE, et al. The Australian Genetics of Depression Study: Study Description and Sample Characteristics. bioRxiv 2019.

24.	Davies MR, Kalsi G, Armour C, et al. The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety. Behav Res Ther 2019;123:103503.

Genetic Risk Prediction of Supreme Cognitive Ability; An Exceptional Case Study

Meike Bartels[1,*], Danielle Posthuma[1,*]

[1] Vrije Universiteit Amsterdam
[*] These authors contributed equally

Intelligence is highly heritable[1] and a major determinant of human health and well-being. While its molecular genetic underpinnings have long remained elusive, the past decade has seen some major breakthroughs. In 2006 a team led by the fantastic and utterly charming dr. Martin reported the discovery of a specific location on the long arm of chromosome 2 to be linked to individual differences in intelligence[2]. In addition, effects of triple repeat polymorphisms in the *SCA1*, *MJD*, and *DPRLA* genes were reported to be associated with several cognitive phenotypes[3]. Recent genome-wide meta-analyses have further identified 24 genomic loci linked to variation in intelligence[4], offering novel insight into the biology of intelligence.

While these studies have shed light on the causes of individual differences in normal cognitive function, they may also be informative for rarer, more extreme forms of intellectual function[5]. One highly understudied condition is *Supreme Cognitive Ability*. This extraordinary rare genetic trait is characterized by extremely lucid bouts of intensely intelligent remarks, and may severely influence a patient's life as well as the patient's direct environment, including family, friends and colleagues, both continental and overseas.

To study the genetic underpinning of this rare genetic trait using insight from normal cognitive function, we here report a case study of an exceptional patient suspected of suffering from *Supreme Cognitive Ability*: Prof. dr. N.G. Martin.

Study Design
Power analyses indicated that a sample size of N=1 was sufficient to detect the presence or absence of genetic risk factors for *Supreme Cognitive Ability*. We decided not to opt for informed consent or approval from the local IRB, but instead decided to leave the subject ignorant of any participation in the current study. To this end, we instructed a pool of younger researchers (**Figure 1**) who needed to gain credits for their PhDs, were not too busy finishing their theses, or were already in a collaborative relationship with Prof. dr. N.G. Martin. These researchers have followed and shadowed the study subject at several locations during the past 5 years. They were, after several failed attempts, able to collect a DNA sample from an obvious object in the subject's life (**Figure 2**).

Fig. 1| Pool of younger researchers tasked to collect DNA from study subject

After careful transport of the used item, DNA of the participant was collected and sent off to an unspecified commercial company. The newest genotyping procedures were applied as well as stringent quality control procedures and after filtering for MAF, HWE, and missingness, a total of 15,789,99 variants were available for further analyses. The unnamed company then selected a small sample of genes that are known to predict the presence of *Supreme Cognitive Ability* as well as several of its correlates and tested these for the presence of risk variants in the subject's genetic code.

Results

Supreme Cognitive Ability: Based on the subject's own remarkable work in the field of intelligence it has long been hypothesized that a person must have two variants of the extremely rare minor allele in the *SCA1* gene in order to have this condition. The sequenced DNA of Prof. dr. N.G. Martin proved to contain two copies of the rare variant of the *SCA1* gene. The company provided us with the following report (**Figure 3**):

Fig. 2 | Wine glass confiscated for DNA sample collection

Nick, you have the two variants in the SCA1 gene

You could still have a variant not covered by this test.

2 variants detected
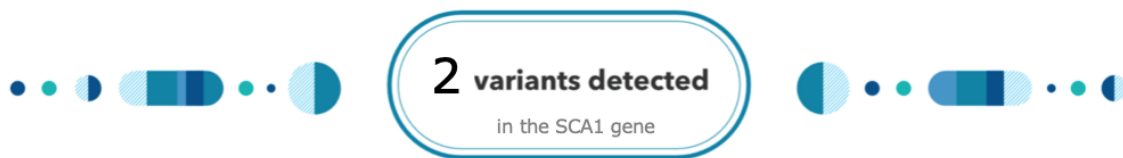in the SCA1 gene

Fig. 3| Report from unnamed company on presence of SCA1 alleles

In addition, the company provided further explanation, to ensure the subject and his direct environment would understand his condition as well as the symptoms that come along with the condition. The following list is advised to be used in 'clinical' practice and maybe even in daily life (**Figure 4**):

About Supreme Cognitive Ability

When symptoms develop
Symptoms typically develop during infancy and get worse and worse throughout the life span

How it's treated
There is currently no known cure. Treatment focuses on managing complications such as uncontrollable urges to start novel projects, and unusual working hours. It could possibly be treated by serving more wine or sending off on holidays

Typical signs and symptoms
- Uncontrollable urge to work
- Contagious enthusiasm
- High level of imagination
- Energizing social meetings
- Seamless combining of work and social events
- Early morning meditation and early evening wine drinking
- Ordering amazing amounts of (mostly delicious) food for the whole table

Ethnicities most affected
This condition is most common in people of European, Ashkenazi Jewish and Hispanic/Latino descent

Read more at

Genetics Home Reference*

GeneReviews*

Mayo Clinic*

Fig. 4| Symptoms and Signs

Speed of Thinking
Finally, leveraging widely publicly available datasets and summary statistics, our team was able to create amazingly reliable polygenic scores for *Speed of Thinking*, an important component of *Supreme cognitive ability*. Although our expectations of this approach for this extremely complex trait were already high, the report surpassed our imagination (**Figure 5**). We found that the risk profile created from summing all risk alleles for *Speed of Processing* showed that Prof. dr. N.G. Martin scored in the top 0.000001% of all individuals tested for this trait. Mendelian randomization analyses showed a unidirectional causal relationship reflecting a direct causal effect of *Speed of Thinking* on *Supreme Cognitive Ability* (P<0.000001)

Our results thus confirm the tested subject suffers from *Supreme Cognitive Ability*, which is likely caused by an early onset of high levels of *Speed of Thinking*. While later in life several of the symptoms of *Speed of Thinking* will be mediated by self-medicated intake of alcohol in the form of white wine (Sauvignon, no Chardonnay), the score is such that symptoms may still be frequently present and affect hundreds of individuals in the vicinity of the subject. We believe many young and older researchers have felt this influence and are continuing to be affected by this. Further research is needed to assess the specific extent of the positive impact Prof. dr. N.G. Martin has had on current and future generations.

We love you Nick, thanks for teaching and inspiring us! *-Meike & Danielle*

Fig. 5| Genetic Risk prediction for Speed of Thinking

References

1.   Wainwright, M. A., Wright, M. J., Geffen, G. M., Luciano, M. & Martin, N. G. The genetic basis of academic achievement on the Queensland Core Skills Test and its shared genetic variance with IQ. *Behav. Genet.* **35**, 133–145 (2005).
2.   Luciano, M. *et al.* Genome-wide scan of IQ finds significant linkage to a quantitative trait locus on 2q. *Behav. Genet.* **36**, 45–55 (2006).
3.   Luciano, M. *et al.* Effects of SCA1, MJD, and DPRLA triplet repeat polymorphisms on cognitive phenotypes in a normal population of adolescent twins. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* **144B**, 95–100 (2007).
4.   Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
5.   Franić, S. *et al.* Intelligence: shared genetic basis between Mendelian disorders and a polygenic trait. *Eur. J. Hum. Genet. EJHG* **23**, 1378–1383 (2015).

Nick Martin as a Mentor – A Perspective

Matthew C. Keller

Everyone in the field of behavioral genetics knows (and has an opinion about) about Nick Martin. Most have collaborated with him, many have been trained by him, and he has influenced almost everyone in the field in one way or another. Nick has an encyclopedic memory of behavioral genetics findings and a thorough understanding of its theoretical foundation. This, along with his exuberance and natural charm, have made him perhaps *the* central nexus in the behavioral genetics community. Nick has been our field's greatest advocate and its greatest facilitator of collaboration, and the field would have been much different, and much diminished, had he not been a part of it.

Like many now in the field, I was brought into the fold of the wider behavioral genetics community by Nick. I met Nick in 2004 at the International Statistical Genetics Workshop in Boulder, Colorado, where he and several of his PhD students were instructors. I was finishing up my PhD in Social Psychology with a Master's degree in Statistics, but didn't know much at all about behavioral genetics or the methodology used to study it. In typical Nick fashion, by the end of our first meeting, had invited me out to visit QIMR for a few months after I graduated and offered to pay me while there. I spent 3 months working in his lab, and it was a turning point in my career: I have considered myself a behavioral geneticist from that point forward and it's a decision I've never really second-guessed. I have returned to visit QIMR many times in the years since and have formed lifelong collaborations and friendships with the people there.

Two interactions with Nick as a mentor stand out as being particularly formative to my scientific career. The first was an "around the water cooler" discussion we had about a recent finding, at the time, that the genetic variant 5-HTTLPR appeared to modify the role that stress played on depression (Caspi et al., 2003). I came, paper in hand, wide-eyed and credulous as people newly in a field often are, to discuss the findings. To my surprise, Nick was highly skeptical, and laid out for the first time to me why the false-positive rate in science can be much higher than the alpha-level of .05- in particular when the prior probability of a hypothesis being true is low—and why this might be especially so in candidate gene research. This was long before the "reproducibility crisis" in behavioral sciences and before it was widely appreciated that many, and in some fields most, scientific findings are false; certainly that was news to me. Five years after that, as an assistant professor at CU Boulder, an enterprising graduate student, Laramie Duncan, came to me with the beginnings of a review paper on candidate gene-by-environment interactions that she had done for a class. Armed with a skepticism inherited from Nick about the approach, and paired with an intelligent and tenacious collaborator, that paper evolved over many iterations into a critical evaluation of the flimsy evidence supporting the enterprise (Duncan & Keller, 2011). By that time, Nick and colleagues had already published several papers attempting— with little success—to replicate previous candidate gene findings in large, highly powered (relative to candidate gene study) samples (Coventry et al., 2010; Gillespie, Whitfield, Williams, Heath, & Martin, 2005; Hansell et al., 2007; Whitfield et al., 2000; Wray et al., 2008). As has become clear in the years since, Nick's skepticism was well-placed. The candidate gene era stands as a cautionary tale about how a field can mislead itself for years and that science can be painfully slow in self-correcting.

The second interaction was when I showed up to Nick's office with some results showing that an earlier finding associating IQ with estimated autozygosity had not replicated in a new sample. Nick could see I was a bit crestfallen about the results. In a compassionate, if somewhat scolding tone, he told me something to the effect of, "We are not in charge of nature. Our job is merely to report what we find as accurately as we can." It feels odd to me now that I should have reacted so, but I felt thunderstruck, as though a weight of worry and future worry suddenly lifted from my shoulders. Having grown up academically in a field and culture where our job as scientists was to "find interesting things," preferably those that support one's pet theory, it was liberating to realize that, no, my job really was just to report what we'd found, as accurately as possible. An idea as simple as that, and I stopped fretting about how

my studies turned out. I freely admit that I continue to root for one outcome or another, but have since realized that if the question is asked well enough, the answer should be interesting regardless of what it is. Certainly that is something to aspire to, even if not always attainable. I have tried to instill this perspective in my own mentees over the years. And so the "vertical transmission," as a behavioral geneticist would put it, of this philosophy of science has passed down, mentor to mentee, across the generations.

Nick is responsible for the scientific starts of many colleagues: Sarah Medland, Dave Evans, Will Coventry, Manuel Ferreira, Brendan Zietsch, Michelle Luciano, Nathan Gillespie, Tim Bates. These are just some of the graduate students and postdocs who worked in Nick's lab and who overlapped with times I was at QIMR, but there are many more who were there before or after. My perspective on his mentorship is but one. Others would talk of his exuberance, his generosity, his willingness to listen, his advocacy of junior researchers, or his impatience with lazy thinkers. But I would hazard to guess that we all would agree that we would not be where we are without him. Nick has helped instill in us a passion for exploration, a healthy skepticism of everyone's findings, including our own, and a sense of duty in trying, at least, to get the answer right, regardless of what that answer is.

References

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., . . . Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science, 301*(5631), 386-389. doi:10.1126/science.1083968

Coventry, W. L., James, M. R., Eaves, L. J., Gordon, S. D., Gillespie, N. A., Ryan, L., . . . Wray, N. R. (2010). Do 5HTTLPR and stress interact in risk for depression and suicidality? Item response analyses of a large sample. *Am J Med Genet B Neuropsychiatr Genet, 153B*(3), 757-765. doi:10.1002/ajmg.b.31044

Duncan, L. E., & Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am J Psychiatry, 168*(10), 1041-1049. doi:10.1176/appi.ajp.2011.11020191

Gillespie, N. A., Whitfield, J. B., Williams, B., Heath, A. C., & Martin, N. G. (2005). The relationship between stressful life events, the serotonin transporter (5-HTTLPR) genotype and major depression. *Psychol Med, 35*(1), 101-111. doi:10.1017/s0033291704002727

Hansell, N. K., James, M. R., Duffy, D. L., Birley, A. J., Luciano, M., Geffen, G. M., . . . Martin, N. G. (2007). Effect of the BDNF V166M polymorphism on working memory in healthy adolescents. *Genes Brain Behav, 6*(3), 260-268. doi:10.1111/j.1601-183X.2006.00254.x

Whitfield, J. B., Pang, D., Bucholz, K. K., Madden, P. A., Heath, A. C., Statham, D. J., & Martin, N. G. (2000). Monoamine oxidase: associations with alcohol dependence, smoking and other measures of psychopathology. *Psychol Med, 30*(2), 443-454. doi:10.1017/s0033291799001798

Wray, N. R., James, M. R., Dumenil, T., Handoko, H. Y., Lind, P. A., Montgomery, G. W., & Martin, N. G. (2008). Association study of candidate variants of COMT with neuroticism, anxiety and depression. *Am J Med Genet B Neuropsychiatr Genet, 147B*(7), 1314-1318. doi:10.1002/ajmg.b.30744

Twin Cohorts

Jaakko Kaprio & Dorret Boomsma

Over the past five to six decades, our knowledge of the etiology of common, complex diseases has deepened enormously. Early epidemiological studies of coronary heart disease in the 1950s such as the Framingham study and the international multisite Seven Countries study identified major risk factors such as smoking, blood lipids and blood pressure. Other large-scale epidemiological studies of cardiovascular diseases, cancer and other common diseases followed over the next decades. In the course of these studies, the importance of family history became evident, and large-scale twin studies were established to enable the distinction between exposures and experiences shared by family members ("shared environment") and genetic factors common to family members.

While the first and nationwide cohorts were established in Nordic countries, (first Denmark, Sweden and then Norway and Finland), large cohorts of twins and their family members have been established in the Netherlands and Australia as well as in many other countries later on. Nick Martin established the Australian Twin Registry in 1978. He has been responsible for the development and expansion of twin and twin-family studies based at QIMR ever since. The importance of twin studies for medical and lots of other traits was emphasized in a key paper that he published in 1997 (Martin et al, 1997). As summarized in this review paper, multiple twin and twin-family studies confirmed the role of genetic factors for nearly every human trait, including those known to be risk factors for common diseases. As relatively rare conditions, twin studies of diseases from any single cohort were generally underpowered to provide reliable estimates of heritability except for the most common conditions. Combining data and analyses from individual cohorts permit more reliable estimates to be made, but also permit analysis of whether variance components differed by country or other aspects of the contributing cohorts, such as an analysis of smoking behavior from Australia, Finland and Sweden (Madden et al, 1999), which then led to the large family-based Nicotine Addiction Genetics study (Saccone et al, 2007). While such ad hoc studies had been done earlier, the GenomEUtwin study was the first large-scale effort to enable a pooled analysis of twin data of complex diseases and their risk factors.

The GenomEUtwin study undertook genome-wide analyses of European twin and population cohorts to identify genes in common diseases and traits, including migraine, BMI, lipids and body height. Data sets were derived from eight twin cohorts from Istituto Superiore di Sanità, Italy, Karolinska Institutet, Sweden, University of Helsinki, Finland, Vrije Universiteit Amsterdam, Netherlands, University of Southern Denmark, Odense, Denmark, Norwegian Institute of Public Health, Norway, St Thomas Hospital, London, UK, and Queensland Institute of Medical Research, Australia as well as other centers contributing expertise in analyses and genotyping. Though funded (13.7 million euros) by the EU through its Framework Programme 5 integrated project funding mechanism, agreements between the EU and Australia allowed Australian researchers to join. QIMR under Nick's leadership was an important component of the project. A first key set of papers were published in the October 2003 issue of Twin Research. The collaborative papers confirmed and extended knowledge of the genetic basis of these traits and became key cited papers. The heritability estimates were very close to each other in all these eight populations of European ancestry, despite quite divergent geographical, cultural and health system environments, for example on migraine (Mulder et al, 2003). Many other analyses followed, and at the end of the project large scale genotyping of MZ pairs using the Illumina 370 chip was conducted to study variability genes in lipid traits (Surakka et al, 2012).

GenomEUtwin laid the foundation for twin cohorts to contribute to the GWAS era of studies. A major and early international effort in this area was the ENGAGE ((European Network for Genetic and Genomic Epidemiology) project. Starting in 2008, the five-year project (www.euengage.org) was a EU-funded large-scale project to enable meta-analyses of genome-wide association studies (GWAS) and develop their methodology. QIMR and Nick Martin have played a major role in developing methods in genetic epidemiology, and he and his group made important contributions to ENGAGE (Aulchenko

et al, 2009). At that time, multiple other consortia for conducting meta-analyses of GWAS data were formed and cohorts that were contributing to ENGAGE were also in many of these other projects. Lessons on how to share summary statistics from individual GWAS analyses, efficiently meta-analyze and interpret results were learnt during these years, and have enabled the stunning success of GWAS studies.

ENGAGE built and expanded on the experience of and accumulated trust in collaborative research. Large-scale collaborations in human genetics have been needed to identify the myriad contributing genes of small effect size in complex disease. The multifactorial, polygenic nature of common traits and diseases was predicted by quantitative genetic theory and empirically seen in twin and family studies. After GenomEUtwin and ENGAGE, Nick has continued to be very active in multiple molecular genetic consortia, several of which he initiated. Examples include the Twinning Genetics Consortium (TGC) (http://www.twinningconsortium.org/), the Psychiatric Genomics Consortium (PGC) (https://www.med.unc.edu/pgc/), Enhancing NeuroImaging Genetics through Meta-Analysis (Enigma) (http://enigma.ini.usc.edu/) and " **ACTION: A**ggression in **C**hildren: Unraveling gene-environment interplay to inform **T**reatment and **I**nterventi**ON** strategies" (Action) (http://www.action-euproject.eu/).

Major human groups of European, African and Asian ancestry have distinct genetic differences, which contribute to differences in genetic risk to common diseases and traits between populations. As large twin studies have been established in many countries, comparisons of heritability estimates and variance components between regions of the world can provide insights into the interplay of genetics and environment. The COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) project (Silventoinen et al, 2019) is a collaborative effort of 54 twin projects from 24 countries, including Australian data from QIMR. Data on weight and height as well as relevant covariates are available on 489,981 twin individuals from both twins from 228,635 twin pairs). Though this is the largest single analysis of twin data to date, Australian data have importantly contributed to many other projects such as the effort to identify genes underlying human twinning – a topic that is dear to Nick.

In addition to these individual projects, Nick has been highly influential in the International Society for Twin Studies, notably as the editor in chief of *Twin Research and Human Genetics*. The journal has published theme issues on twin cohorts and registries in 2002, 2006, 2013 and in 2020. These document the scope and impact of twin research and research about and of twins and other multiples, to which Nick has made important contributions.

References

Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, Penninx BW, Janssens AC, Wilson JF, Spector T, Martin NG, Pedersen NL, Kyvik KO, Kaprio J, Hofman A, Freimer NB, Jarvelin MR, Gyllensten U, Campbell H, Rudan I, Johansson A, Marroni F, Hayward C, Vitart V, Jonasson I, Pattaro C, Wright A, Hastie N, Pichler I, Hicks AA, Falchi M, Willemsen G, Hottenga JJ, de Geus EJ, Montgomery GW, Whitfield J, Magnusson P, Saharinen J, Perola M, Silander K, Isaacs A, Sijbrands EJ, Uitterlinden AG, Witteman JC, Oostra BA, Elliott P, Ruokonen A, Sabatti C, Gieger C, Meitinger T, Kronenberg F, Döring A, Wichmann HE, Smit JH, McCarthy MI, van Duijn CM, Peltonen L; ENGAGE Consortium. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet*. 2009;41:47-55.

Madden PA, Heath AC, Pedersen NL, Kaprio J, Koskenvuo MJ, Martin NG. The genetics of smoking persistence in men and women: a multicultural study. *Behav Genet*. 1999;29:423-31.

Martin N, Boomsma D, Machin G. A twin-pronged attack on complex traits. *Nat Genet*. 1997;17:387-92.

Mulder EJ, Van Baal C, Gaist D, Kallela M, Kaprio J, Svensson DA, Nyholt DR, Martin NG, MacGregor *Twin Res* AJ, Cherkas LF, Boomsma DI, Palotie A. Genetic and environmental influences on migraine: a twin study across six countries. . 2003;6:422-3

Saccone SF, Pergadia ML, Loukola A, Broms U, Montgomery GW, Wang JC, Agrawal A, Dick DM, Heath AC, Todorov AA, Maunu H, Heikkila K, Morley KI, Rice JP, Todd RD, Kaprio J, Peltonen L, Martin NG, Goate AM, Madden PA. Genetic linkage to chromosome 22q12 for a heavy-smoking quantitative trait in two independent samples. *Am J Hum Genet*. 2007;80:856-66.

Silventoinen K, Jelenkovic A, Yokoyama Y, Sund R, Sugawara M, Tanaka M, Matsumoto S, Bogl LH, Freitas DL, Maia JA, Hjelmborg JVB, Aaltonen S, Piirtola M, Latvala A, Calais-Ferreira L, Oliveira VC, Ferreira PH, Ji F, Ning F, Pang Z, Ordoñana JR, Sánchez-Romera JF, Colodro-Conde L, Burt SA, Klump KL, Martin NG, Medland SE, Montgomery GW, Kandler C, McAdams TA, Eley TC, Gregory AM, Saudino KJ, Dubois L, Boivin M, Brendgen M, Dionne G, Vitaro F, Tarnoki AD, Tarnoki DL, Haworth CMA, Plomin R, Öncel SY, Aliev F, Medda E, Nisticò L, Toccaceli V, Craig JM, Saffery R, Siribaddana SH, Hotopf M, Sumathipala A, Rijsdijk F, Jeong HU, Spector T, Mangino M, Lachance G, Gatz M, Butler DA, Gao W, Yu C, Li L, Bayasgalan G, Narandalai D, Harden KP, Tucker-Drob EM, Christensen K, Skytthe A, Kyvik KO, Derom CA, Vlietinck RF, Loos RJF, Cozen W, Hwang AE, Mack TM, He M, Ding X, Silberg JL, Maes HH, Cutler TL, Hopper JL, Magnusson PKE, Pedersen NL, Dahl Aslan AK, Baker LA, Tuvblad C, Bjerregaard-Andersen M, Beck-Nielsen H, Sodemann M, Ullemar V, Almqvist C, Tan Q, Zhang D, Swan GE, Krasnow R, Jang KL, Knafo-Noam A, Mankuta D, Abramson L, Lichtenstein P, Krueger RF, McGue M, Pahlen S, Tynelius P, Rasmussen F, Duncan GE, Buchwald D, Corley RP, Huibregtse BM, Nelson TL, Whitfield KE, Franz CE, Kremen WS, Lyons MJ, Ooki S, Brandt I, Nilsen TS, Harris JR, Sung J, Park HA, Lee J, Lee SJ, Willemsen G, Bartels M, van Beijsterveldt CEM, Llewellyn CH, Fisher A, Rebato E, Busjahn A, Tomizawa R, Inui F, Watanabe M, Honda C, Sakai N, Hur YM, Sørensen TIA, Boomsma DI, Kaprio J. The CODATwins Project: The Current Status and Recent Findings of Collaborative Project of Development of Anthropometrical Measures in Twins. *Twin Res Hum Genet*. 2019; 31:1-9.

Surakka I, Whitfield JB, Perola M, Visscher PM, Montgomery GW, Falchi M, Willemsen G, de Geus EJ, Magnusson PK, Christensen K, Sørensen TI, Pietiläinen KH, Rantanen T, Silander K, Widén E, Muilu J, Rahman I, Liljedahl U, Syvänen AC, Palotie A, Kaprio J, Kyvik KO, Pedersen NL, Boomsma DI, Spector T, Martin NG, Ripatti S, Peltonen L; GenomEUtwin Project. A genome-wide association study of monozygotic twin-pairs suggests a locus related to variability of serum high-density lipoprotein cholesterol. *Twin Res Hum Genet*. 2012;15:691-9.

Human Sexuality

Karin Verweij and Brendan Zietsch

Nick Martin does not fear taking risks, he does not bend to criticism, he can be bold, and he doesn't strive to be politically correct. This may have made him some enemies over his career, but it has also won him admiration and respect and has stood him far out from the crowd.

Unlike many other scientists, Nick does not shun controversial research topics. In 1992, he – with Michael Bailey – developed a questionnaire to send out to twins to investigate the underpinnings of sexuality. Before then, little genetic research had been done on human sexuality, probably in part because of squeamishness and decorum on the part of the scientific community. That didn't hold Nick back. The 12-page questionnaire contained general items about the participants' background, personality, family composition, handedness etcetera, but also very personal questions about their sexual behaviour, including items about sexual orientation, sociosexuality, and sexual behaviour (*e.g. "How many times have you done the following?: Sexual intercourse, entering vagina from rear"*). Despite the sensitive nature of the questionnaire, it was completed by almost 5000 twins. To ensure participants' anonymity, thus minimising the threat of discomfort and dishonest answers, Nick and Michael came up with a solution whereby the twins were asked to agree on a 10 digit number which they would each enter on the top of their questionnaires. In this way, questionnaires could not be related to individual twins but twin pairs could be matched up. (Unfortunately, this also meant that the data could not be matched with their genotype data years later). Of all twins that were asked to participate, 27% explicitly refused, 19% initially agreed but subsequently did not return the consent forms, and 52% explicitly consented[1]. There were some small differences between twins who participated and those who did not, suggesting that results of the questionnaire may slightly overestimate sexual liberalism, activity, and adversity[1]. Those who responded had higher education levels, scored higher on novelty seeking, were less conservative (both on voting preferences and on sexual attitudes), and attended church less often. Responders also showed higher prevalence of depression, alcohol dependence, conduct disorder, and reported an earlier age at sexual intercourse, and higher rates of sexual abuse[1]. While 27% of twins who were approached may have been put off by the invitation, the actual participation rate of more than 50% was not bad at all.

Once the data were collected, they were used to investigate many interesting research questions, and this is where we (Brendan and Karin) come in. When, in 2007, Karin wanted to do a research internship abroad for her master's degree, Dorret Boomsma suggested she go to Brisbane to work with Nick. Initially Karin wanted to work on something to do with substance use, but Nick convinced her to work on the sex questionnaire that hadn't been used much. Karin was happy to work on a topic that was a bit more distinctive, so 'sex and genetics' it was, and she never regretted that choice - just looking at the descriptive statistics was already interesting. In the meantime, Brendan had been plugging away at a multivariate twin model of EEG data for almost a year, when Karin told him about the existence of the sex questionnaire. For Brendan this questionnaire triggered his genuine interest in research and launched a career in which he applied genetics methodology to questions about human mating and how it relates to the evolution of human nature.

With the data from the sex questionnaire Nick, Brendan, Karin, and others explored genetic influences on various interesting traits. This yielded the first clear evidence of heritability of

sexual orientation[2], the female orgasm[3], homophobia[4], sociosexuality[5], and risky sexual behaviour[6,7].

We also used the twin data to look into potential evolutionary explanations of the maintenance of homosexuality in the population, which many see as a Darwinian Paradox[8]. We found evidence that psychologically masculine females and feminine men are more likely to be nonheterosexual, but when heterosexual, they have more opposite-sex sexual partners. We showed that these relationships are partly due to genetic influences common to each trait. We also find a trend for heterosexuals with a nonheterosexual twin to have more opposite-sex partners than do heterosexual twin pairs. These results suggest that genes predisposing to homosexuality may confer a mating advantage in heterosexuals, which could help explain the evolution and maintenance of homosexuality in the population. Notably, at the moment we are re-investigating this hypothesis using genome-wide genotype data.

In two other studies we looked at the association between sexual orientation and personality and mental health[9,10]. Previous research indicated that homosexuals and bisexuals are, on average, at greater risk for psychiatric problems than heterosexuals, potentially because of prejudice often experienced by nonheterosexuals. We tested whether apparent sexual orientation differences in psychiatric vulnerability simply mirrored sex differences in personality traits, i.e. nonheterosexual males having elevated neuroticism scores as females do, and nonheterosexual females having elevated psychoticism scores as males do. Our results contradicted this, with nonheterosexual men and women scoring significantly higher on both neuroticism and psychoticism than their heterosexual counterparts, suggesting an overall elevation of psychiatric risk in nonheterosexuals (as neuroticism and psychoticism are positively associated with psychiatric disorders). We also found significant genetic correlation of sexual orientation with neuroticism and psychoticism, but no corresponding environmental correlations[9]. Similarly, in a subsequent paper[10] we showed that non-heterosexual men and women had elevated rates of lifetime depression and that genetic factors accounted for a majority (60%) of the correlation between sexual orientation and depression. In addition, childhood sexual abuse and risky family environment were significant predictors of both sexual orientation and depression, further contributing to their correlation. These findings do not mean that anti-gay prejudice has no effect on psychiatric vulnerability in non-heterosexuals, but they do suggest there is more to the story.

For a paper on testing evolutionary theories of the female orgasm, we worked with renowned evolutionist Geoffrey Miller, whom Nick recruited to the department for a sabbatical. Geoffrey's book The Mating Mind was one of Brendan's inspirations for getting into evolutionary psychology, so it was a fantastic opportunity to work together. We found that the female orgasm data didn't fit any of the existing adaptive theories, leaving us stuck with the question: Why does it exist? Like many of the most interesting questions, we still don't know the answer – despite several subsequent papers from Brendan using other approaches and datasets – but with Nick's help and support, these questions have at least been raised to the status of deserving serious inquiry with serious data, and we're sure as the answers come in it will be in no small part thanks to Nick.

On a broader level, Nick has obviously meant a lot to the field of behavioural genetics. His pioneering papers, mentorship, vision of how science should be performed, large data collection and generosity in sharing these data, as well as his encouragement of collaboration and open science have played a defining role in the field. On a personal level, Nick has meant a lot to Karin's and Brendan's careers. He encouraged us to pursue our genuine interests, and

from the start of our PhD he gave us the freedom to develop our own research questions and approaches. He stimulated us to work hard, to collaborate with others, and to not shy away from the hard questions. The way he mentored us was stimulating and greatly contributed to us becoming independent scientists. Brendan, Karin, and Nick still share the interest in investigating biological factors in human sexuality and research questions that can be off-road. Brendan and Karin both still very happily collaborate with Nick on various projects. Nick's continuing motivation and dedication to keep doing research and his sincere interest in genetics are unparalleled and he we hope he will keep up the good work for many more years to come.

References

1   Dunne, M. P. *et al.* Participation bias in a sexuality survey: psychological and behavioural characteristics of responders and non-responders. *International journal of epidemiology* **26**, 844-854, doi:10.1093/ije/26.4.844 (1997).

2   Bailey, J. M., Dunne, M. P. & Martin, N. G. Genetic and environmental influences on sexual orientation and its correlates in an Australian twin sample. *Journal of personality and social psychology* **78**, 524-536, doi:10.1037/0022-3514.78.3.524 (2000).

3   Dawood, K., Kirk, K. M., Bailey, J. M., Andrews, P. W. & Martin, N. G. Genetic and environmental influences on the frequency of orgasm in women. *Twin Research and Human Genetics* **8**, 27-33, doi:10.1375/1832427053435427 (2005).

4   Verweij, K. J. *et al.* Genetic and environmental influences on individual differences in attitudes toward homosexuality: an Australian twin study. *Behav Genet* **38**, 257-265, doi:10.1007/s10519-008-9200-9 (2008).

5   Bailey, J. M., Kirk, K. M., Zhu, G., Dunne, M. P. & Martin, N. G. Do individual differences in sociosexuality represent genetic or environmentally contingent strategies? Evidence from the Australian twin registry. *Journal of personality and social psychology* **78**, 537-545, doi:10.1037/0022-3514.78.3.537 (2000).

6   Zietsch, B. P., Verweij, K. J. H., Bailey, J. M., Wright, M. J. & Martin, N. G. Genetic and environmental influences on risky sexual behaviour and its relationship with personality. *Behavior Genetics* **40**, 12-21, doi:10.1007/s10519-009-9300-1 (2010).

7   Verweij, K. J. H., Zietsch, B. P., Bailey, J. M. & Martin, N. G. Shared aetiology of risky sexual behaviour and adolescent misconduct: genetic and environmental influences. *Genes Brain and Behavior* **8**, 107-113, doi:10.1111/j.1601-183X.2008.00456.x (2009).

8   Zietsch, B. P. *et al.* Genetic factors predisposing to homosexuality may increase mating success in heterosexuals. *Evolution and Human Behavior* **29**, 424-433, doi:10.1016/j.evolhumbehav.2009.07.002 (2008).

9   Zietsch, B. P., Verweij, K. J. H., Bailey, J. M., Wright, M. J. & Martin, N. G. Sexual Orientation and Psychiatric Vulnerability: A Twin Study of Neuroticism and Psychoticism. *Archives of sexual behavior* **40**, 133-142, doi:10.1007/s10508-009-9508-4 (2011).

10   Zietsch, B. P. *et al.* Do shared etiological factors contribute to the relationship between sexual orientation and depression? *Psychol Med* **42**, 521-532, doi:10.1017/s0033291711001577 (2012).

(Helsinki, 1995)

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM (2010. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*, 42(7), 565-9. (cites = 3259)

SNP-based Heritability – a Commentary on Yang et al. (2010)

Jian Yang

Before I moved to the field of human genetics, I was working on quantitative trait locus (QTL) mapping in experimental populations of plants and animals. That is why I did not know the name Nick Martin until the second last year of my PhD candidature. Towards the end of my PhD, it was clear to me that I should find a postdoc position somewhere, but Australia was not quite on my radar until the year 2006 when I had a 3-month visit to Western Australia. Then, I started thinking about the possibility of moving to Australia. A few Google searches brought my attention to the research groups led by Professors Nicholas Martin, Grant Montgomery, and Peter Visscher. I joined Peter's lab in September 2008 to start my academic career in human genetics.

I might have seen Nick at my job interview seminar at the Queensland Institute of Medical Research (renamed QIMR Berghofer Medical Research Institute in 2013), but to my memory, we met for the first time when Peter introduced me to him in his office. The conversation was short but impressive not least because I saw the old fashioned computer on his desk. I was also impressed later on when I often saw him working in the office on a Sunday afternoon, which, believe me, is not common in Australia.

In the year 2009, I was working with Peter (and Mike Goddard from Melbourne) on a project aiming to estimate the proportion of variance in human height explained by all single nucleotide polymorphisms (SNPs) that are common in European populations. At that time, there was confusion about the genetic architecture of common traits and diseases like height and obesity largely because of the observation that genetic loci identified from published genome-wide association studies (GWASs) only accounted for a small fraction of heritability for almost all the traits studied, leading to the "missing heritability" puzzle and criticisms of the failure of GWAS as an experimental design.

In GWAS, each SNP is tested for association with a trait of interest one by one across the genome to search for genomic loci responsible for the trait variation in a population. Because of the large number of tests performed (typically from 100,000s to millions depending on the coverage of the SNP array and whether the SNP data have been imputed to a reference panel with whole-genome sequence data), a correction for multiple testing is needed to avoid false-positive discoveries, e.g., a p-value threshold of 5e-8 is often used to claim significant findings from GWASs. This means that if the effect size of a SNP is small and the GWAS sample size is not sufficiently large, we would not have enough power to detect it at a genome-wide significance level. Hence, a critical question was how much proportion of the trait variance is accounted for by the SNPs that did not reach genome-wide significance. This might be achieved by fitting the effects of all SNPs jointly as random effects in a mixed linear model.

The model was appealing, but how about the data? It was not like these days when we can easily get access to GWAS data sets of 10,000s or even 100,000s of individuals from public resources such as dbGaP and the UK Biobank. GWASs with only a few thousand or even a hundred individuals were not common at that time. Our model attempted to estimate the aggregated effect of many SNPs, which is equivalent to a classical additive genetic model $y = g + e$ with g being the total additive genetic value of an individual captured by all SNPs and e being the residual. Estimating the variance of g and thereby the heritability captured by all SNPs, i.e., the SNP-based heritability $h^2_{SNP} = var(g) / var(y)$, required a correlation matrix of g (also known as the genetic relationship matrix or GRM). We did not want to include any related individuals in the model because otherwise, we could not distinguish whether the estimated var(g) was captured by the SNPs or by the pedigree relatedness reconstructed from the SNP data. The latter is more complex and can contain variance components due to common environmental effects that are shared among close relatives and rare genetic variations not tagged by array SNPs. The precision of the estimate of var(g) (often measured by the standard error or SE), however, is inversely proportional to the variability of the off-diagonal elements of the GRM. Because the model uses only unrelated individuals, the variance of the off-diagonal elements of the GRM is small so that a relatively large sample size (at least much larger than those used in pedigree-based heritability analyses) is required to obtain an estimate of h^2_{SNP} with useful precision.

We started with an analysis of a data set with ~2500 unrelated people and an estimate of h^2_{SNP} for height that was somewhere between 0.4 and 0.5. We were all very excited about it, but the SE and thus the confidence interval of the estimate was too wide to make any convincing conclusion. Fortunately, we heard from Nick that there was an additional batch of data that would be available soon, which pushed the sample size up to ~4000. We finally obtained an estimate of 0.45 (SE = 0.08), which was significantly larger than the proportion of variance accounted for by SNPs passing genome-wide significance (~10%) reported by a GWAS meta-analysis of ~180,000 individuals in 2010.

The implication of this study was profound. It suggested that a large proportion of the heritability for height could be explained by all common SNPs on an array, so that the heritability was not missing but rather hiding in the form of many variants of small effect scattered across the genome. Genome-wide association studies at that time were not very successful mainly because most complex traits are the result of many genetic variants each with an effect too small to reach the stringent genome-wide significance threshold. This suggested that the genetic architecture for height (and possibly for many other common traits and diseases) was likely to be polygenic and that more associations would be discovered in GWASs with larger sample sizes. These findings and their implications have been corroborated by a number of studies in recent years. The paper on this work, entitled "Common SNPs explain a large proportion of the heritability for human height", was eventually published in *Nature Genetics* in 2010 and has received >3000 citations in the past 10 years.

This study would have not been possible without the critical contribution from Nick. The amazing human genetic resources established by the team led by Nick and the critical mass of researchers in human genetics in Brisbane directly and indirectly because of him had laid the foundation for scientific ideas like this to evolve and to be implemented. His generosity in data sharing and vision in human genetics have always inspired me.

The Barbarians are at the Gate!

Pete Hatemi

Why care about attitudes, voting, religion, or politics when humanity faces so many problems closer to our mortality? Indeed, for those who are attempting to cure diseases, treat cancer, help mitigate the onset of schizophrenia, or treat any of those conditions that disrupt the lives of so many people, asking and answering questions about more basic human behaviors might seem less important. In a world of finite resources, exploring the sources of political attitudes and beliefs may appear an endeavor where time and money is perhaps best not spent. Nick's work helped many others see this differently. On the one hand, it might appear more practical to overlook politics and focus on more immediate health concerns. But there has arguably been nothing more devastating to the human species than humans. Politics affects everyone. Attitudes and beliefs in the aggregate shape the world we live in, the rules of society, and how resources are allocated; they regulate the rights, freedoms, and liberties we enjoy, or are denied access to. The wars that continue to be fought over identity, culture, politics and religion; the drive of consumption; the need to have power and control over others; and the devastation that comes from declarations of "us" versus "them", leading to mass suffering, genocides, holocausts, and displacement of peoples - are all the result of political choices. These choices lead to an untold number of deaths, health-related disparities, malnutrition, abuse, stress, depression, deprivation, anxiety, and violence.

While the discipline of political science has been mostly preoccupied with addressing such dilemmas, the field, until recently, had a somewhat odd view of human behavior; mainly that it was absent the human part. The discipline remained largely wedded to a Durkheimian ideal, embedded in a paradigm where social forces and external stimuli constituted the only meaningfully cause of variation in human behavior. Political ideals were nothing more than social constructions - too recent a phenomenon and too context dependent to be passed down through genetic transmission. The blank slate was as real in political science in 2005 as it was for John Watson's "behaviorist manifesto" in 1913. This view was not limited to scholars or academics. Rather, those who regularly influence policy at the highest levels of government- secretary of states, national security advisors, presidents and secretary generals of the UN such as Woodrow Wilson, Zbigniew Brzezinski, Condoleezza Rice, Tijjani Muhammad Bande and Henry Kissinger are only a few; all were trained with such a background.

This myopic view of human behavior radically changed in the mid 2000's, due in large part to the influence of Nick Martin and Lindon Eaves. Nick has led or been a major contributor to scores of studies on the genetic influences of attitudes and ideologies, vote choice, political sophistication, partisan identification, political trust, immigration and out-group attitudes, political violence, morality, economic behaviors, educational attainment, sex differences, threat sensitivity, disgust, risk taking, fitness, fear, aggression, pursuit of power, and rational action, among many other topics (For only a handful of his papers in this area, see Alford et al., 2011; Eaves et al., 2011; Hatemi, Alford, et al., 2009; Hatemi, Funk, et al., 2009; Hatemi et al., 2010; Hatemi et al., 2014; Hatemi et al., 2007; Hatemi et al., 2015; Rietveld et al., 2013; Smith et al., 2017; Sturgis et al., 2010; Verweij et al., 2008; Zietsch et al., 2011). No less than 10 special issues in social science journals have been devoted to this area of research in the last decade. These studies did more than provide mere estimates of heritability, but rather used a wide variety of methods, including gene-environment interaction approaches, assortative mating corrections, longitudinal models, cross-cultural and direction of causation models, extended kinships, and genome wide approaches.

Nick's leadership and scholarship led to a shift in theory as well – genetic influences were not simply operating on political attitudes. In other words, there is no gene for views on gay rights for example, but rather modern attitudes were seen to reflect the same fundamental issues of survival and reproduction that confronted ancient humans. Essentially, modern views on immigration tap into the same emotional and cognitive mechanism surrounding the need to identify and address out-groups. Issues underlying universal health care are modern instantiations of how to share resources; issues of marriage and sex roles are contemporary forms of finding a mate and raising children; and defense and punishment policies, no matter how complex, find their roots in protecting our families and group. The modern manifestation of genetic influences on these traits are complicated by institutions, technology, nation-states and other social movements. Certainly, the labels differ across time and space, but the underlying connection between the core issues of human survival - sex, group identity, food, shelter and defense, remain the same. And some combination of migration, genetic drift, assortative mating, mutation, recombination, culture, life events, and local ecological adaptation drives variation on such traits.

Nick has been involved in some form, in every major twin and genetic study of politics since the 1980s, served as PI on a number of grants from the National Science Foundation (NSF) to study these traits and was a guiding mentor, along with Lindon Eaves in several NSF grants to train political scientists in genetic methods. More than 50 political and social scientists were trained at the IBG/BGA methods workshop through these grants, leading to hundreds of publications that helped transform the field.

In 2008, I went to political science's largest conference and there was a plenary speaker who, almost shouting, declared "We must stop this introduction of genetics into our discipline". He went on, "Can't you see the barbarians are at the gate!" As I looked around at this stadium sized conference hall, most of the audience was shaking their heads in agreement. Now, more than a decade later, in 2020, it is a regular occurrence to find neuroscience, hormones, genetics, and biobehavioral models as mainstream political science. I was just invited to chair a dedicated panel on biobehavioral approaches at this year's conference. It is difficult today to find a discussion on human behaviors, beliefs, conflicts, identity and war, without at least some inclusion of both social and inherited mechanisms.

This transformation would not have occurred without Nick Martin. And this speaks to the importance Nick has played in the field, not only scientifically but also personally, as a mentor. For it is not only science that matters, but the scientist. And in this instance, the best way to explain Nick's role is to describe a bit about my own personal experience with Nick. Nick's influence was not nearly as direct as one might expect or anywhere near within the timeframe expected. Indeed, the foundational works of Eaves, Eysenck and Martin that identified genetic influences on individual differences in political values (Eaves & Eysenck, 1974; Eaves, Eysenck, & Martin, 1989; Martin et al., 1986) went entirely unnoticed for 20 years in the discipline that was most in need of their research. This changed in 2004. By chance, one political scientist, John Hibbing, a congressional scholar at the University of Nebraska came across Lindon's and Nick's work, and through a collaborator at Virginia Commonwealth University, Carolyn Funk, made contact with the good Dr. Eaves. Lindon, gracious as ever, gave Funk access to his data, having no idea what he was about to unleash. Through a series of Monty Python-esque missteps and a different understanding of disciplinary norms, they published on Lindon's data… without Lindon. On a personal level, that was somewhat of a disaster, but the result of that unfortunate misunderstanding led to two incredible happenings. While Hibbing and company's findings were nothing new to geneticists, as they simply re-

presented the results from Martin et al. (1986) and Eaves et al. (1989), it served as the first, albeit accidental, step into introducing an entire discipline to a very different way of thinking.

The second benefit was far more personal. Their mis-step led to my introduction to Nick, although again not in a manner planned or expected. While convalescing with my brother Jon during his time at Law School at the University of Nebraska, he encouraged me to seek a less dangerous profession, and go back to school. Having no clue on what I would study, but knowing a fair bit about violence, conflict and war, I walked into the political science department, and met with Kevin Smith, the graduate director at the time. Thus began the first step in my education with Nick. After a short period, I came to realize two things: political science had the most interesting questions, but it was missing half of the tools to study them. There were remarkably few scholars in political science who considered the importance of inherent differences in cognition, motivation, perception, and attitude development to explain variation in behavior. I was fortunate that by coincidence, fate, or accident, two of those scholars, John Hibbing and Kevin Smith, were at Nebraska. Even more serendipitous is that I came upon Nick's work, independently, having no idea of the interactions they had with Lindon or Nick's relationship to Lindon. After reading Lindon and Nick's work, I had an "a ha" moment: there are inherent individual differences in political and social values that are genetically transmitted across generations. Perhaps this approach offered a way to help answer questions of: why is there a Hitler, or a Pol Pot? Why are some people motivated to engage while others are not? Why are some so ready and able to rise up and fight for an identity that has nothing to do with their personal lives? Why are some willing to kill simply for a label? Why do some seek to elevate others, while some only want to further their own interests? Why do some resist equality while others embrace violence? And what makes us different? Nick and Lindon's initial contribution pointed toward a way of answering these questions that social learning models had yet to reasonably answer. Maybe not in my lifetime, but carrying forward their work was an exciting prospect, and so becoming aware of Hibbing's write-up of Lindon's data, I reached out to Lindon, having no idea what my colleagues at Nebraska had mistakenly done. Lindon's response to me was to "go bugger off". And so, I walked right from one minefield into another. I emailed maybe another 40 -50 scholars in this area wanting to learn how to conduct behavior genetic analyses, where to start, how to collect data and so forth. Almost none responded. Two emailed back with a list of their publications and nothing more. In short, I was repeatedly given nothing but cold shoulders. One person meaningfully responded: Nick Martin, the director of the largest genetic epidemiology lab in the Southern Hemisphere, emailed me back, a part time PhD student with nothing more to offer than my interest. He invited me to come work with him at QIMR. With no training in genetics, no understanding of matrix algebra, structural models, or any real skills in research, I got on a plane and flew out to Australia with not even a place to stay, and began working with Nick Martin. Nick treated me as one of his Ph.D. students, and more than that. He introduced me to Will Coventry, who after a day put me up with a place to stay. And then began my real education in science. It was that email that began and set my career as an academic and a passion for science I never knew I had. And through Nick I met Sarah Medland who must be recognized because this progress in the discipline and my own would not even remotely have been possible without her guidance and help. And it was Nick who mended fences, and contacted Lindon and reminded him, not to blame the son "for the sins of the father." From there, I brought what I learned from Nick's lab to VIPBG and worked with Lindon and Mike Neale, Hermine Maas, Matt Keller and others who pushed the discipline further.

Nick does not just build science, he builds scientists. I tell this story not because my role is important, in truth it is not; but rather because my story is not unique at all when it comes to

Nick. At any one time, you will find a handful of scholars, not Nick's declared students, but people he invests in, simply for their own sake and that of their ideas. The unique element is Nick. Anyone who knows anything about him knows he will support and mentor folks from any country, any place, any background. The only thing they have to do is have an idea and be willing to work it. And there is no metric one can easily point to, to identify the depth of his mentorship. Because unlike most scholars, where there is a defined and official advisor/advisee role, Nick has been an advisor to scores of people who are nowhere on paper associated with him. That is the truest form of mentorship and selfless science. I am honored to write on behalf of all those students to celebrate a part of Nick's contribution that many do not know, and to express the deepest appreciation for what he has done and continues to do for so many who stand on his shoulders and follow in his footsteps. There are so few people in science, or in any industry, like him. In this way, the apple does not fall far from the tree. There are many great scholars, many great intellectuals, many great leaders in the field. Nick is all those. But it is rare when you find someone as committed to the science as to the scientist. I would not be an academic, an intellectual, or a scholar today without Nick's guidance, leadership, friendship, mentorship, and care. The field of political science would most likely still be living in behaviorism. And I have no doubt there are scores of other people who would say the exact same thing about their path in science. Thank you, Nick.

References

Alford, J. R., Hatemi, P. K., Hibbing, J. R., Martin, N. G., & Eaves, L. J. (2011). The Politics of Mate Choice. *The Journal of Politics, 73*(02), 362-379.

Eaves, L. J., & Eysenck, H. J. (1974). Genetics and the development of social attitudes. *Nature, 249*(454), 288-289.

Eaves, L. J., Eysenck, H. J., & Martin, N. G. (1989). *Genes, culture, and personality : an empirical approach*. London ; San Diego: Academic Press.

Eaves, L. J., Hatemi, P. K., Heath, A. C., & Martin, N. G. (2011). Modeling Biological and Cultural Inheritance. In P. K. Hatemi & R. McDermott (Eds.), *Man is by nature a political animal evolution, biology, and politics* (pp. 101-184). Chicago: University of Chicago Press.

Hatemi, P. K., Alford, J. R., Hibbing, J. R., Martin, N. G., & Eaves, L. J. (2009). Is There a "Party" in Your Genes? *Political Research Quarterly, 62*(3), 584-600.

Hatemi, P. K., Funk, C. L., Medland, S. E., Maes, H. M., Silberg, J. L., Martin, N. G., & Eaves, L. J. (2009). Genetic and Environmental Transmission of Political Attitudes Over a Life Time. *The Journal of Politics, 71*(03), 1141-1156.

Hatemi, P. K., Hibbing, J. R., Medland, S. E., Keller, M. C., Alford, J. R., Smith, K. B., . . . Eaves, L. J. (2010). Not by twins alone: Using the extended family design to investigate genetic influence on political beliefs. *American Journal of Politcal Science, 54*(3), 798-814.

Hatemi, P. K., Medland, S. E., Klemmensen, R., Oskarsson, S., Littvay, L., Dawes, C. T., . . . Martin, N. G. (2014). Genetic Influences on Political Ideologies: Twin Analyses of 19 Measures of Political Ideologies from Five Democracies and Genome-Wide Findings from Three Populations. *Behavior Genetics, 44*(3), 282-294.

Hatemi, P. K., Medland, S. E., Morley, K. I., Heath, A. C., & Martin, N. G. (2007). The genetics of voting: an Australian twin study. *Behavior Genetics, 37*(3), 435-448.

Hatemi, P. K., Smith, K., Alford, J. R., Martin, N. G., & Hibbing, J. R. (2015). The genetic and environmental foundations of political, psychological, social, and economic behaviors: a panel study of twins and families. *Twin research and human genetics, 18*(3), 243-255.

Martin, N. G., Eaves, L. J., Heath, A. C., Jardine, R., Feingold, L. M., & Eysenck, H. J. (1986). Transmission of social attitudes. *Proceedings of the National Academy of Sciences, USA, 83*(12), 4364-4368.

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., . . . Koellinger, P. D. (2013). GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science, 340*(6139), 1467-1471.

Smith, K., Alford, J. R., Hibbing, J. R., Martin, N. G., & Hatemi, P. K. (2017). Intuitive Ethics and Political Orientations: Testing Moral Foundations as a Theory of Political Ideology. *American Journal of Politcal Science, 61*(2), 424–437.

Sturgis, P., Read, S., Hatemi, P., Zhu, G., Trull, T., Wright, M., & Martin, N. (2010). A Genetic Basis for Social Trust? *Political Behavior, 32*(2), 205-230.

Verweij, K. J., Shekar, S. N., Zietsch, B. P., Eaves, L. J., Bailey, J. M., Boomsma, D. I., & Martin, N. G. (2008). Genetic and environmental influences on individual differences in attitudes toward homosexuality: An Australian twin study. *Behavior genetics, 38*(3), 257-265.

Zietsch, B. P., Verweij, K. J., Heath, A. C., & Martin, N. G. (2011). Variation in human mate choice: simultaneously investigating heritability, parental influence, sexual imprinting, and assortative mating. *The American Naturalist, 177*(5), 605-616.

(Early Boulder Faculty)

Hansen JL, Reed DR, Wright MJ, Martin NG, Breslin PAS (2006). Heritability and Genetic Covariation of Sensitivity to PROP, SOA, Quinine HCl, and Caffeine. *Chemical Senses*, 31(5), 403–413. (cites = 84)

Hwang LD, Zhu G, Breslin PA, Reed DR, Martin NG, Wright MJ (2015). A common genetic influence on human intensity ratings of sugars and high-potency sweeteners. Twin Res Hum Genet, 18(4), 361-7. (cites = 15)

Hwang LD, Breslin P, Reed D, Zhu G, Martin NG, Wright MJ (2016). Is the Association Between Sweet and Bitter Perception due to Genetic Variation? *Chem Senses*. 1;41(9):737-744. (cites = 14)

Do People with Lower IQ Have Weaker Taste Perception? A Hidden Supplementary Table in "Is the Association Between Sweet and Bitter Perception due to Genetic Variation?

Daniel Hwang

This paper is about how a study on the genetic association between the perception of sweetness and bitterness ended up looking at the influence of intelligence on taste.

The story started from a chat between Nick Martin and Danielle Reed at the 2001 American Society of Human Genetic Conference. The two scientists had mutual interest in taste perception and decided to build an international collaboration, together with Margie Wright and Paul Breslin, to collect taste and smell data from twins in Australia and the USA to explore the human genetics of chemical senses.

The first study (Hensen et al., 2006) from this collaboration quantified the heritability of the perceived intensity of four bitter substances — propylthiouracil ($h^2 = 0.72$), caffeine ($h^2 = 0.30$), quinine ($h^2 = 0.34$) and sucrose octaacetate ($h^2 = 0.28$). The paper brought quantitative genetics to chemosensory sciences and is one of the most highly cited papers in *Chemical Senses* (84 cites).

I joined the collaboration in 2010 as a research technician working with Danielle at the Monell Chemical Senses Center (Philadelphia, USA). He was responsible for collecting sensory data at the annual Twins Day Festival in a city named Twinsburg in Ohio. He also prepared the taste solutions and shipped them to Australia for taste tests.

In 2014 I moved to Brisbane to do my PhD with Nick and Margie Wright at the QIMR. I learned how to use the classic MX software to perform structural equation modelling from Gu Zhu for analyzing twin data. My first project (Hwang et al., 2015) established the heritability of the perception of sugars and artificial sweeteners ($h^2 = 0.30 — 0.34$) and identified a common genetic factor.

It had been known for more than a decade that human sweet and bitter taste receptors were both G protein-coupled receptors; however, no one had ever proved a shared genetic mechanism. This question became the second project of my PhD which aimed to investigate the genetic correlation between sweet and bitter taste perception.

In August 2015, I presented results from bivariate and multivariate modelling to Nick and Margie. At the meeting Nick was suspicious about whether the moderate genetic correlation ($r_g = 0.46 — 0.51$) was due to confounding and hypothesized that people with lower IQ might have weaker taste perception. *"Why don't you have a look at IQ in our 19UP study?"*, Nick said to me. This suggestion gave IQ a place in the paper.

Surprisingly, IQ was correlated with taste perception but in the opposite direction to what Nick thought. People with lower IQ actually rated both sweet and bitter solutions more intense (Supplementary Table 8). Nevertheless, including IQ (as well as the Big 5 personality traits) did not change the genetic correlation between the perception of sweet and bitter tastes.

Supplementary Table 8. Phenotypic correlations between taste intensities and IQ, personality and emphasis scores estimated from bivariate ACE models

	IQ	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness	Emphasis
PROP	-0.11*	0.04	0.02	-0.05	-0.07*	-0.03	-0.02
SOA	-0.15*	0.07*	0.03	-0.07*+	-0.06*+	-0.04	-0.02
Quinine	-0.14*	0.07*	0.05	-0.05	-0.04	-0.05	0
Caffeine	-0.13*	0.07*	0.02	-0.04	-0.06*+	-0.04	-0.02
gSweet	-0.07*	0.05	0.02	0.00	-0.03	-0.05	0

n = 1244~1256. *p < 0.05 before correction for multiple testing. +Insignificant after adjusting for IQ.

The paper was published in 2016 and later recommended in F1000 due to its significance in solving a decade-long question in chemosensory sciences. However, the association between IQ and taste was only briefly discussed as "higher IQ is associated with less extreme rating styles" and hidden in the supplementary materials, and whether intelligence does influence our taste perception remains a puzzle.

We still don't know if people who always complain about foods are too sweet or too bitter have different intelligence levels than others, but it may not necessarily be a bad thing to be a non-taster!

Sociopolitical Attitudes Through the Lens of Behavioral Genetics: Contributions from Dr. Nicholas Martin

Brad Verhulst

The idea that political attitudes are heritable remains a contested hypothesis in political science. When Dr. Nicholas Martin first published his seminal paper on the heritability of attitudes in the Proceedings of the National Academy of Sciences nearly 35 years ago (Martin et al. 1986), the proposition was unfathomable to most social scientists. Then, along with Drs. Eaves and Eysenck, Dr. Martin expanded his inquiry into the transmission of social attitudes, replicating his findings and pairing the investigation with the transmission of personality traits. Still, this research program was generally ignored in Political Science which would have been the most natural audience given the subject matter, however psychologists, especially those within the field of behavioral genetics made occasional references to it (e.g. Bouchard et al., 1990; Tesser et al., 1993). The Political Science community finally took notice of the possibility that political attitudes were heritable after the publication of Alford Funk and Hibbing's (2005) manuscript that recapitulated the decades-old findings. Even 20 years after his initial publication, the suggestion that attitudes could be anything other than socially constructed was treated as heresy. In their fury, opponents of the proposition attacked the assumptions of the twin model (Charney, 2008; Suhay, 2007) or questioned whether it was ethical to explore possible genetic components of attitudes. At academic conferences, scientists presenting results about the heritability of attitudes were accused of being eugenicists. As the attacks raged, Dr. Martin counseled patience. He had witnessed similar inquisitions of twin methods in the 1970s and 1980s and with the confidence drawn from previous experience understood that behavioral genetic methods would again prevail over the detractors.

Social Attitudes as a Model Phenotype

While the heritability of political attitudes was ignored by political scientists, within behavioral genetics it was treated as a model phenotype due to its unique mode of intergenerational transmission. Specifically, political attitudes have a significant additive genetic component, a significant shared environmental component, and a significant unshared environmental component. By contrast, most adult psychiatric disorders and psychological behaviors tend to be characterized by additive genetic and unique environmental components (childhood and adolescent behaviors occasionally have a significant shared environmental component, the importance of which decreases at older ages). Furthermore, there is a substantial spousal correlation for political attitudes. Because of the mode of intergenerational transmission, political attitudes were often treated as a model phenotype to test a variety of methodological components of the classical twin design.

Using attitudes as model phenotypes, Dr. Martin was able to explore two essential assumptions of the twin model to demonstrate its robustness: violations of the equal environment assumption and violations of the random mating assumption. These assumptions are necessary to obtain unbiased estimates of the parameters from twin models.

The equal environments assumption requires that the environments of MZ and DZ twins are functionally equivalent: that the environments MZ twins find themselves in are not more similar than the environments of DZ twins. Opponents of the twin model often opine that because MZ twins are more genetically similar than DZ twins, the world treats them more similarly than DZ twins. A cursory perusal of the items that contribute to classical zygosity assessments appears to suggest that parents (a central focus of children's environment) unwittingly treat MZ twins more similar than DZ twins. For example, as children MZ twins are more likely to share the same bedroom, be dressed in matching clothing, and have the same friends, relative to their DZ counter parts, and this environmental similarity at early ages could lead to enhanced similarity later in life. Thus, the quintessential assumption is that MZ twins are more similar because they spend more time together. This assumption, however, can be empirically tested with longitudinal twin data. What Dr. Martin and his colleagues found was that attitudinal similarity leads to social contact, and reciprocally, that contact leads to increased similarity (Posner et al., 1996). As such, it is not the case that environmental treatment drives twins to be more

similar, but instead that internal motivations, driven in part by genetic factors, leads MZ twins to phenotypic similarity and social contact. This clearly contradicts the expectations regarding what would be observed if violations of the equal environment assumption were driving MZ similarity.

Another place where Dr. Martin was able to leverage the unique components of social attitudes was to explore the implications of assortative mating on heritability. The random mating assumption of the classical twin model requires that spouses are uncorrelated for the traits of interest. While there is minimal spousal resemblance between most personality traits or psychiatric disorders, for social attitudes and other social phenotypes such as religiosity there is a substantial spousal correlation (Zietsch et al., 2012). Spousal phenotypic correlations increase the genetic similarity between DZ twins for the particular phenotype, consequently increasing the DZ phenotypic correlation. If assortative mating is ignored, it can inflate the estimates of the shared environment and deflate the estimate of the heritability.

When considered jointly, the lack of violation of the equal environments assumption paired with the failure of the random assortment assumption imply that the heritability of political attitudes is actually larger than what would be expected by a standard twin model.

Direct Contributions to Political Science

Dr. Martin has co-authored more manuscripts that have appeared in top tier political science journals than most card-carrying political scientists, but perhaps his greatest contribution to political science was the prescience of adding items that assess social attitudes to his own and other twin studies. These include the Canberra Twins study (circa 1980s), the extended study of Australian Twins (circa ~ 1989-1994), and the Virginia 30,000 Study of Twins and their relatives (circa ~ 1988-1994). The addition of political attitudes items to these studies made it feasible to explore the possibility that social attitudes had a genetic component and has laid the foundations for mapping the biology of these traits. With these data in hand, Dr. Martin encouraged collaborations where he would generously provide guidance and mentorship to scholars around the globe who were interested in exploring these questions (but his generosity extended well beyond the field of political science). His collaborations with political scientists has fueled a resurgence of research in modes of transmission of cultural values and a long overdue paradigm change in the field.

References

Baker, L.A., Treloar, S.A., Reynolds, C.A., Heath, A.C., & Martin, N.G. (1996) Genetics of Educational Attainment in Australian Twins: Sex Differences and Secular Changes. *Behavior Genetics*, 26(2); 89-102.

Bouchard, T.J. Jr., Lykyen, D.T., McGue, M., Segal, N.L., Tellegen, A. (1990) Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart. *Science* 250

Charney, E. (2008) Genes and Ideologies. *Perspectives on Politics*, 6(2): 299-319.

Martin, N.G., Eaves, L.J., Heath, A.C., Jardine, R., Feingold, L.M., Eysenck, H.J. (1986). Transmission of Social Attitudes. *Proceedings of the National Academy of Sciences of the United States of America*, 83(12): 4364-4368.

Posner, S.F., Baker, L., Heath, A., Martin, N.G., (1996) Social contact, social attitudes, and twin similarity. *Behavior Genetics* 26 (2), 123-133.

Suhay, E, Kalmoe, N., & McDermott, C. (2007) Why twin studies are problematic for the study of political ideology: Rethinking "Are Political Orientations Genetically Transmitted?" Paper presented at the annual meeting of the International Society of Political Psychology, Portland, OR.

Tesser, A. (1993) The Importance of Heritability in Psychological Research: The case of attitudes. *Psychological Review*, 100(1): 129-142

Zietsch, B.P., Verweij, K.J.H., Heath, A.C., & Martin, N.G. (2012) Variation in human mate choice: Simultaneously investigating heritability, parental influence, sexual imprinting, and assortative mating. *The American Naturalist*, 177(5): 605-616.

(Nick, Georgia and Nick's parents)

Lind, P.A., Zhu, G., Montgomery, G.W., Madden, P.A.F., Heath, A.C., Martin, N.G., & Slutske, W.S. (2013). Genome-wide association study of a quantitative disordered gambling trait. *Addiction Biology, 18*, 511-522. (cites = 53)

Slutske, W.S., Zhu, G., Meier, M.H., & Martin, N.G. (2010). Genetic and environmental influences on disordered gambling in men and women. *Archives of General Psychiatry, 67*, 624-630. (cites = 110)

Gambling Disorder

Wendy S. Slutske and Penelope A. Lind

Nick Martin is the least likely person to set foot in a casino or purchase a lottery ticket. He is too sensible and is eager to remark that "the lottery is a tax on the foolish." Yet, it was mostly good luck that led to his becoming one of the world leaders in the effort to discover the genetic underpinnings of disordered gambling.

Nick and his colleagues have been conducting ground-breaking investigations of the genetics of alcohol use disorder based in his Genetic Epidemiology Unit at QIMR for nearly four decades, and we were fortunate to be a part of these efforts. When one of us (WS) learned that the highest per-capita spending on gambling in the world was in Australia, she felt virtually obligated (!) to study disordered gambling in the Australian Twin Registry (ATR). When Nick was approached with idea, he was all in. We were able to obtain funding from the National Institute of Mental Health (NIMH; one of the few funded projects that focused on gambling disorder before NIMH decided that gambling disorder was a low priority) to conduct a survey of gambling behaviors and disorder in the ATR. Data from the Australian Twin Study of Gambling (OZ-GAM) have been a gold mine for new discoveries about gambling and gambling disorder – there are no other data like these in the world. To date, 17 gambling-related publications have been based on these data.

This represented only the second large-scale twin study of disordered gambling ever conducted, and the only one that had included women. As expected, we were able to demonstrate that disordered gambling was heritable, that it was equally heritable in men and women, and that the association between disordered gambling and alcohol use disorder was largely due to genetic factors. These data also allowed us to confirm or refute established wisdom from the social sciences about the role of the environment in disordered gambling. For example, among discordant twin pairs (that is, after controlling for genetic and shared environmental factors), the twin who spent more time gambling with her parents was no more likely to frequently gamble or to develop gambling problems than the twin who spent less time gambling with her parents. And among discordant pairs, the twin who began to gamble at a younger age was no more likely to go on to develop gambling problems than the twin who started to gamble at an older age. On the other hand, among discordant pairs, the twin who lived in a more disadvantaged neighborhood was more likely to develop gambling problems than the twin who lived in a more advantaged neighborhood.

Relatively few genetic association studies of disordered gambling have been conducted; all but two have been candidate gene studies and the majority of those have focused on genes in the dopaminergic system. It was not until 2013 that we published the first ever gambling-related genome-wide association study (GWAS) with Nick in *Addiction Biology*; the power of the study was limited by the low number of genotyped OZ-GAM participants (n=1,312) and no genome-wide significant SNPs were identified. However, we reported three novel loci for disordered gambling with highly suggestive evidence of association and enriched biological pathways that were previously associated with substance addiction. We then contributed data to the only other published gambling GWAS which was also underpowered to identify genome-wide significant loci.

While WS was visiting QIMR, Nick drew to our attention a report from the Australian Adverse Drug Reactions Bulletin for the drug cabergoline, a potent dopamine D_2 receptor agonist. There were reports of four patients taking long-term levodopa for the treatment of Parkinson's disease who began to gamble excessively a few months after cabergoline was added, and whose gambling problems abated when the cabergoline was discontinued. Similar adverse drug effects were reported in North America with the dopamine receptor agonist pramipexole. These pharmacologic findings provided a compelling clue to a potential neurobiological pathway to disordered gambling that might prove to be useful in gene identification. To explore this further we conducted a gene-enrichment analysis within the OZ-GAM study using a gene set derived from the literature on dopamine-induced disordered gambling and candidate gene studies of gambling disorder; we observed enrichment of association with disordered gambling at both the level of the SNP and the gene.

Lack of federal funding for gambling disorder research has slowed down but has not blocked efforts to move forward. In the ensuing years, Nick, PL and colleagues have continued to deliberately include measures of disordered gambling (and of course, GWAS genotyping) in new data collection

projects at QIMR, including two new twin samples and large national efforts to recruit individuals with histories of depression or bipolar disorder. Polygenic risk scores for bipolar disorder and Parkinson's disease were created in two of these samples as predictors of disordered gambling. There was a significant association between the genetic risk for Parkinson's disease and disordered gambling, but not between genetic risk for bipolar disorder and disordered gambling. This was the first study to demonstrate a genetic link between Parkinson's disease and disordered gambling, and is especially intriguing in light of the pharmacologic findings described above. There is now converging evidence from two distinct lines of inquiry suggesting that the pathophysiology underlying Parkinson's disease may play an important role in the etiology of disordered gambling.

Genetic risk variants for disordered gambling have not yet been robustly identified due in part to the limited availability, at an international level, of existing community and clinically ascertained cohorts with DNA samples. Mostly due to Nick's foresight, we are now moving in the direction of having an adequately-sized GWAS meta-analysis of disordered gambling by including our QIMR-based Australian studies and cohorts from the United Kingdom, Germany, and the United States -- what we are optimistically calling "GD1," with the expectation that it will be the first in a series of gambling disorder GWAS meta-analyses. Establishment of this international consortium was important as it is only through large collaborative GWAS meta-analyses that we will get a clearer picture of genetic mechanisms contributing to disordered gambling. For the first time we will be able to examine whether genetic risk for disordered gambling overlaps with genetic contributions to psychiatric and medical comorbidities, as well as educational attainment, cognitive functioning, brain structure, and personality traits.

From a genetic perspective, gambling disorder has been an "orphan disorder." It has been the subject of very few twin or genomic studies, and is one of the few major mental disorders that has not been included in the Psychiatric Genomics Consortium. Looking to the future, we are optimistic that many of the great insights about the etiology of disordered gambling will come from our fledgling international GWAS consortium, which would not exist without the intellectual curiosity and generosity of Nick Martin.

References

Davis, C.N., Slutske, W.S., Martin, N.G., Agrawal, A., & Lynskey, M.T. (2019). Genetic and environmental influences on gambling disorder liability: A replication and combined analysis of two twin studies. *Psychological Medicine, 49*, 1705-1712.

Isaacs, D., Lander, C., McNeil, J., Pillans, P., Strasser, S., & Wainwright, D. (2005). Australian adverse drug reactions bulletin. *Australian Adverse Drug Reactions Bulletin, 24(2)*.

Lang M., Leménager T., Streit F., Fauth-Bühler M., Frank J., Juraeva D., Witt S.H., Degenhardt F., Hofmann A., Heilmann-Heimbach S., Kiefer F., Brors B., Grabe H.J., John U., Bischof A., Bischof G., Völker U., Homuth G., Beutel M., Lind P.A., Medland. S.E., Slutske W.S., Martin N.G., Völzke H., Nöthen M.M., Meyer C., Rumpf H.J., Wurst F.M., Rietschel M., Mann K.F. (2016). Genome-wide association study of pathological gambling. *European Psychiatry*, *36*, 38-46

Lind, P.A., Campos, A., Colodro-Conde, L., Medland, S.E., Slutske, W.S., & Martin, N.G. (October, 2019). Prediction of pathological gambling and problematic gambling behaviours by bipolar disorder and Parkinson's disease: A polygenic risk score analysis. Poster presented at the annual meeting of the American Society of Human Genetics, Houston, TX.

Lind, P.A., Zhu, G., Montgomery, G.W., Madden, P.A.F., Heath, A.C., Martin, N.G., & Slutske, W.S. (2013). Genome-wide association study of a quantitative disordered gambling trait. *Addiction Biology, 18*, 511-522.

Slutske, W.S., Deutsch, A.R., Richmond-Rakerd, L.S., Chernyavskiy, P., Statham, D.J., & Martin, N.G. (2014). Test of a potential causal influence of earlier age of gambling initiation on gambling involvement and disorder: A multi-level discordant twin design. *Psychology of Addictive Behaviors, 28*, 1177-1189.

Slutske, W.S., Piasecki, T.M., Deutsch, A.R., Statham, D.B., & Martin, N.G. (2019). Potential causal influence of neighborhood disadvantage on disordered gambling: Evidence from a multi-level discordant twin design. *Clinical Psychological Science, 7*, 582-596.

Slutske, W.S., Piasecki, T.M., Ellingson, J.M., & Martin, N.G. (2010). The family history method in disordered gambling research: A comparison of reports from discordant twin pairs. *Twin Research and Human Genetics, 13*, 340-346.

Slutske, W.S., Zhu, G., Meier, M.H., & Martin, N.G. (2010). Genetic and environmental influences on disordered gambling in men and women. *Archives of General Psychiatry, 67*, 624-630.

Cannabis Research

Eske Derks, Karin Verweij, Nathan Gillespie

The International Cannabis Consortium (ICC) was founded in 2013 by Jacqueline Vink, Nathan Gillespie, Karin Verweij and Eske Derks. The first meta-analysis of this consortium was published in the journal *Translational Psychiatry*. This meta-analysis included 32,330 subjects (discovery) + 5,627 (replication). We did not identify any genome-wide significant SNPs but found four genes to be associated with lifetime cannabis use. The Supplementary Table shows the sample sizes from the 13 individual cohorts, these range from 338 to 6,778. Not surprisingly, the largest contribution came from Nick, who shared data from two cohorts, together comprising 7,499 samples (23% of the total discovery sample).

You may think "who cares about these numbers? It's about the science and the results". This is not actually true since there would be no science without these numbers. In fact, Nick has made enormous contributions to GWA studies of a very wide range of traits and disorders and often was the one contributing the largest sample to a meta-analysis. These days, it may be easy to forget how much time and energy was spend on collecting these samples. As a member of the genetics community, I will always be grateful for Nick's contribution to science as I would not have been able to do any of my work, without the contributions of Nick and others (e.g., my PhD supervisor Prof. Dorret Boomsma who also collected very large sample sizes).

Since our first publication, we have been able to further increase sample sizes and the most recent study of the ICC was published in the influential journal *Nature Neuroscience*. This study has revealed important new insights into the genetic risk factors of cannabis use (most notably the influence of the gene *CADM2* which also impacts other traits, such as risk seeking behaviour and impulsivity). We were also able to investigate the hypothesis that cannabis use leads to schizophrenia. This hypothesis has been generally accepted by the public and the scientific community, but as we all know, correlation does not imply causation. Using Mendelian randomization analysis, we were able to show that in contrast to what many believe, individuals at higher genetic risk for schizophrenia have a higher chance of cannabis use initiation. This suggests that the relation between cannabis use and schizophrenia is more complex than proposed by others, as schizophrenia liability seems to be a precursor of cannabis use. We did not find any evidence for cannabis use causing schizophrenia, but were cautious to make a strong conclusion about a complete lack of causal association in this direction as we had relatively low statistical power to explore this direction of effect. The relationship may still be bidirectional. Our results do show that individuals at higher genetic risk for schizophrenia are more likely to use cannabis, possibly as a form of self-medication.

Our findings are just one example of how genetic data allows investigation of causal relations which are often found to be in unexpected directions. Nick seems to love to challenge old ideas and is unafraid to voice his opposition or propose alternative explanations. Indeed, a print of the following quote by Hans Eysenck is shown in Nick's office and states "*I always felt that a scientist owes the world only one thing, and that is the truth as he sees it. If the truth contradicts deeply held beliefs, that is too bad. Tact and diplomacy are fine in international relations, in politics, perhaps even in business; in science only one thing matters, and that is the facts*".

We have talked about how Nick has collected large sample sizes, and we have mentioned science. But we have not yet mentioned Nick's influence on junior researchers by supporting their development and growth into senior positions. This can be exemplified by the ICC which

is led by four senior investigators. Of these four researchers, three are directly connected to Nick. Karin came to Nick's group first in 2007 for a 6 months internship, and then again in 2008 to start her PhD under Nick's supervision. A major part of her PhD was on the genetics of cannabis use. She enjoyed this period very much and proceeded with this line of research in future positions. Last year she was appointed professor at the Amsterdam UMC and Nick flew over from Brisbane to watch her inaugural lecture entitled "Drugs and Genes", showing their strong scientific and personal bond. Eske started at QIMR Berghofer in 2017 and would not have been in Brisbane without Nick's support. His generosity has given her ample opportunities to set up new studies and publish in influential journals. Nathan completed his PhD under Nick Martin's supervision in 2004 before moving to Richmond, VA where he has worked at the Virginia Institute for Psychiatric and Behavior Genetics (VIPBG) ever since. Incidentally, Nick also undertook post-doctoral training at VIPBG before returning to Australia to establish the Australian Twin Registry. Despite the distance, Nathan has maintained a very active collaboration with Nick over the years and secured NIH funding to collect cannabis data on the Brisbane Longitudinal Twin Study that have been used as part of the ICC meta-analyses.

Following-the advent and increased application of genome wide association scan studies two decades ago, Nick Martin foresaw the need for international consortia as the best means of increasing power to detect alleles for complex behaviours and disorders. He was instrumental in encouraging us to pool our expertise and resources to establish the ICC in 2013. We are very grateful to have followed his advice.

In summary, Nick Martin has influenced today's science by collecting large and genetically informative data, encouraging and participating in large international consortia, and promoting open science. Nick has not only facilitated our ability to begin unravelling the aetiology of so many complex disorders, but in doing so, he has encouraged and fostered many junior researchers to develop themselves into independent researchers and respected leaders of their fields of expertise. It would be premature to call this collection of letters a Festschrift. Nick's motivation and dedication to behavioural genetics is unparalleled and shows no signs of slowing down. Given his recent grant success, Nick will continue his pioneering and unprecedented work and we look forward to many more years of collaboration.

Nick Martin- A Personal Note

Eske Derks

I first met Nick as a PhD student in the early 2000's, when I attended the twin workshops in Boulder, meetings of the Behavior Genetics Association and the twin workshops in Egmond, the Netherlands. As a junior scientist, I was very impressed by Nick's and my PhD supervisor, Dorret Boomsma's enormous achievements. They both initiated large twin and family registers which formed the basis of many heritability and genome-wide association studies.

Fast forward to early 2016 when I worked as professor of Genetics in Psychiatry at the Academic Medical Center in Amsterdam…. I received multiple emails from colleagues who made me aware of an open position as group leader at QIMR Berghofer. After spending 6 months in Brisbane in 2010/2011, I was quite interested and the first step for me to take was to email Nick to discuss my potential application for this position. Nick was very enthusiastic and supportive and he gave me and my family a warm welcome when I came over for a job interview. My decision to make the move down under was driven by two main reasons. First, Brisbane is one of the best places to be for quantitative genetics thanks to Nick's huge impact on the field and his ability to "raise" and attract many brilliant researchers. Second, as a family we fell in love with the beautiful nature and beaches of Queensland. Admittedly, Nick does not have a direct influence on this but he definitely has chosen an amazing place to spend his career.

It is difficult to overstate Nick's impact on the field. He initiated and led the collection of many large-scaled studies on every trait or disorder you can think of. Nick wants these data to be used. Period. He asks nothing in return for this, but is driven by a large curiosity to learn the answers to novel and interesting scientific questions. His generosity has opened avenues for researchers around the globe and has provided them with opportunities to develop themselves as independent researchers. It is too easy to forget about all the work that Nick has done to generate resources that form the basis of genetics research today. Finally, he has an amazing ability to predict which topics will receive interest from the scientific community (and he will never hesitate to tell you what he thinks).

Back to the science… Nick has influenced behavioural (and non-behavioural) science by exploring genetic influences on any trait we can think of. This includes intelligence, hair structure, handedness, psychiatric diseases, and substance use (disorders). When I showed him our results on the latest cannabis GWAS meta-analysis (published in *Nature Neuroscience*) and he noticed the strong genetic correlation with risk-taking behaviour, Nick commented "Of course that makes sense, no person who is conscientious, like I am, would decide to use cannabis. I have never used cannabis".

Nick, I would like to thank you for being so generous and for being a living example of "Open Science". I'm sure you will do more great things in the next five years (or more)! It will be a pleasure to work with you!

(Lindon, Nick and Georgia, Wedding 1983)

Colodro-Conde, L., B. Couvy-Duchesne, John B. Whitfield, Fabian Streit, Scott Gordon, Kathryn E. Kemper, Loic Yengo, et al. 2018. Association Between Population Density and Genetic Risk for Schizophrenia. *JAMA Psychiatry* 75 (9): 901–10. (cites = 13)

Nick Martin's Contribution to GxE Research

Lucía Colodro-Conde, Baptiste Couvy-Duchesne

Nick Martin's interest in genotype-environment interactions (GxE) and genotype-environment correlation (covGE) can be traced back to his PhD thesis in Birmingham (N. G. Martin 1976). Part of his PhD work on the topic was published in the 1977 paper "*A progressive approach to non-additivity and genotype-environmental covariance in the analysis of human differences*" (Eaves et al. 1977), co-authored with Lindon J. Eaves (Nick's PhD advisor), Krystyna A. Last and John L. Jinks (Lindon's PhD advisor and precursor on the topic (Jinks and Fulker 1970)). The article of 42 pages reviews the interpretation, estimation and statistical power of many difficult concepts (including GxE, covGE, assortative mating). Together with Dorret Boomsma (Boomsma and Martin 2002), we find the first statement of the abstract still accurate more than 40 years later, despite many subsequent publications on these topics.

> "*No aspect of human behaviour genetics has caused more confusion and generated more obscurantism than the analysis and interpretation of the various types of non-additivity and non independence of gene and environmental action and genotype-environment interaction and covariation, dominance and assortative mating*" **Eaves et al., 1977**, *p.1.*

In this article, Nick Martin is credited for empirically demonstrating during his PhD that interactions may be dependent on the choice of scale (N. G. Martin 1976; Eaves et al. 1977). This is particularly important for psychometric scales (such as the personality and attitude factors used as examples) dependent on item selection, item weighting, and scale transformation. Thus, scales must be chosen for their interpretability and that of the resulting statistics, keeping in mind there is not such a thing as a "true" scale. In addition, even if a change of scale may represent a change of trait, sensitivity analyses may be used to evaluate the effect of scale distribution on the conclusions (Eaves et al. 1977).

In addition, Nick suggested the presence of GxE on behavioural traits (N. G. Martin 1976) by studying MZ pairs raised together (Jinks and Fulker 1970). This design relies on the fact that any GxE effect introduces a correlation between the MZ pair mean and absolute intra-pair variance. Thus, a correlation between absolute within pair differences and the mean value of an MZ pair suggests the presence of GxE, although it may also point towards an interaction between shared and unique environmental sources of variance (Jinks and Fulker 1970). To note, estimating the GxE variance components (GxC and GxE) is limited by the fact it requires an extended twin design, with twins reared together and apart, as well as unrelated individuals reared together (Eaves et al. 1977; Jinks and Fulker 1970). However, unmodelled GxE can bias the heritability and environmental estimates from twin models (Eaves et al. 1977; Jinks and Fulker 1970).

In 1987, Nick, Lindon Eaves and Andrew Heath performed simulations to estimate the statistical power of GxE analyses that use measured genetic loci and environmental risk factors (N. G. Martin, Eaves, and Heath 1987). The authors considered an ascertained twin design, which estimated the main effects and the interaction of the measured genotype and environment and controlled for background genetic and environmental sources of variation. In addition, it allowed estimation of epistasis (interaction between measured loci and background genetics) as well as between measured environment and background genetics. This model was visionary, in that it prefigured controlling for background genetics (i.e. population structure) in association testing, while introducing genetic interaction analyses. Take it genome-wide and

you may recognise a modern linear mixed model genome-wide association study (GWAS; (Yang et al. 2014) or a genome-wide environment interaction study (GWEIS, Dunn et al. 2016). In addition, the article reports the important increase in statistical power arising from studying a continuous (cancer liability) over a discrete (cancer diagnosis) phenotype, which relates to the discussion on scale we alluded to previously.

A decade later, Andrew Heath, Lindon Eaves and Nick published the results of a twin model for a depression score, stratified by marital status, and concluded in favour of a modifying effect on the genetic liability for depression (Heath, Eaves, and Martin 1998). In addition to his work on behaviour, psychology, and psychiatry, Nick also contributed to other areas of medical research such as skin cancer - of special interest for Queensland, which displays one of the highest prevalence in the world (Staples et al. 2006; "Cancer in Australia: An Overview 2012, Table of Contents - Australian Institute of Health and Welfare" n.d.). A GxE investigation in 2002 examined the association between sun exposure and skin cancer, stratifying the analyses by familial risk (Siskind et al. 2002). The authors concluded there was an interaction between familial risk (proxy for the cancer genetic liability) and sun exposure, although it was not directly tested.

It should come as no surprise that Nick was invited to contribute a book chapter about GxE concepts and methods (Boomsma and Martin 2002; D'Haenen 2002), a very well written and documented introduction to GxE. Another review article, focused on GxE in the context of alcohol use and twin models (Heath et al. 2002). It reiterates limitations in power of estimating the GxE variance components, or the need for twin designs to include twins reared apart. It also envisages that linkage analyses could pinpoint relevant loci, which would offer a direct measurement of genetic liability:

> *Analyses of genotype x environment interaction effects will always be more powerful when genotypes as well as environments can be measured. In the alcohol field, the identification of polymorphisms that affect alcohol metabolism that are associated with differences in alcohol dependence risk offers rich, although as yet underexploited, opportunities for studying such effects.* **Heath et al., 2002, p.35**.

When GWAS started to identify robust and replicated genetic loci, for example the *FTO* variant associated with BMI (Cornes et al. 2009), Nick's GxE investigations resumed, this time focusing on individual genetic variants. In this case, the authors compared the intra-pair variance of MZ twins for each *FTO* SNP status, and also tested the interaction between *FTO* variants and parity in women- although none were significant (Cornes et al. 2009).

It is worth mentioning another article co-authored by Nick which studies the genetic contributions shared between socio-economic status (SES) and gambling (Slutske et al. 2015). By using the GxE twin model proposed by Shaun Purcell (Purcell 2002), the authors showed a significant increase in genetic and environmental variance in gambling as a function of SES. In addition, the article reported that SES, often thought to be an environmental exposure, had a genetic component and showed a genetic correlation with gambling behaviour (Slutske et al. 2015). In contrast, a similar analysis on the genetics of IQ failed to identify a significant interaction with SES (Bates et al. 2016).

We have focussed on articles where Nick is first or last author, but a quick search of his bibliography returns at least another 14 publications relating to GxE that he has contributed to.

In 2015, and more than 1,000 papers after his PhD, Nick had a project to propose to us – to directly test the diathesis-stress model for the origins of depression. At this time, we were two very early career scientists (working in our PhD and first year postdoc respectively). At the time, psychiatric polygenic risk scores (PRS) were starting to show some level of prediction (Wray, Goddard, and Visscher 2007). We embraced Nick's project with enthusiasm: it meant a great opportunity to continue our work together, to learn and practice statistical skills, and to empirically test one of the main theories for the origins of depression. In our innocence we did not foresee that the project would take us more than two years of hard work to complete. It was however totally worth the effort and "the diathesis-stress project" is to date one of our main scientific accomplishments (Colodro-Conde et al. 2017). Two similar initiatives were conducted at the time by other groups, which considered childhood trauma and stressful life events as environmental exposures (Musliner et al. 2015; Peyrot et al. 2014). Previous articles on the topic had used candidate gene approach, where the interaction was tested for with a single gene or a handful of loci (sometimes not robustly associated), leading to inconsistent results.

The "diathesis-stress" design benefited from the recent availability of predictive polygenic risk scores, a direct measure of the diathesis (i.e the genetic vulnerability/predisposition for a trait), in our case, depression. This made it possible to test for GxE using observed G and E (Heath et al. 2002). In practice, we tested the association between a depression score and the diathesis for depression (approximated by PRS, "G"), stressful life events (stress score, "E") and their interaction (GxE). In such a model, the GxE interaction captures the multiplicative effect of genetic predisposition and environmental exposures on top of their additive contribution to the risk of depression. Data for the study was already available thanks to previous data collections (by Nick and colleagues) (Kirk et al. 2000; Gillespie et al. 1999; Treloar et al. 1999).

The diathesis-stress study fits nicely in Nick's body of work. As previously flagged (Heath et al. 2002) this approach offered additional power compared to a variance component analysis where the G and/or E factors are not specified (Jinks and Fulker 1970; Purcell 2002). In addition, although limited to the genetic liability tagged by the GWAS summary statistics and the list of stressors collected, the GxE effect benefits from a greater interpretability (compared to a global variance components), the sign of the interaction being one example (Eaves et al. 1977). Finally, this GxE investigation built on results from robust GWAS and methodological developments relative to genetic risk prediction, as anticipated in previous publications (Heath et al. 2002).

Our "diathesis-stress" meetings took place every Tuesday. It was the three of us plus Gu Zhu, and Sarah Medland. Gu had worked with the stress and depression data and provided the Item Response Theory (IRT) score variables, while Sarah (who was Lucía's supervisor) contributed her statistical expertise and critical thoughts along the whole project. Nick told us many times (and experience proved him right) that regular meetings are the only way to get a project moving. At one point, Nick reached out to other GxE experts to validate our approach and results. You can probably guess who he called: Andrew Heath and Lindon Eaves.

As we had undertaken this project in addition to our other workloads, we necessarily had to work some weekends. We usually met on Saturday morning at the markets, before eating together and working on the paper for the rest of the day, which sometimes extended to the Sunday (often with some party in between). Being a person who enjoys devoting his Sundays to work, read, and catch up with the literature, you could tell Nick was very proud of us and eager to hear every Monday about our studious weekend (as well as about the party). The truth is that we all enjoyed (with some doses of pain) every step of the process. This includes

clarifying concepts, the formulation of hypotheses, the computation of every variable and design of the analyses, as well as the huge amount of checks that we performed to convince ourselves the results were real. Nick took on every opportunity to make us think and actively participate in all discussions – and so we did. We also appreciated his unlimited memory of the data collected at QIMR or of the specific tables or figures in the papers that he wanted us to cite.

Nick arranged for us to have early access to the unpublished summary statistics of the GWAS meta-analysis run by the Psychiatric Genomics Consortium. We showed (Colodro-Conde et al. 2017) that the PRS computed with the updated GWAS offered a stronger measure of the diathesis for depression than the GWAS (PGC-MDD1, Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al. 2013) used in previous publications (Musliner et al. 2015; Peyrot et al. 2014). More importantly, we found a significant positive interaction between the PRS for depression and the scores of (personal) stressful life events, accounting for social support and network life events, and controlling for the population structure (twin sample). This effect was replicated about a year later in the Generation Scotland dataset (Arnau-Soler et al. 2019).

However, the project did not finish there and through a challenging revision process we kept building on what Nick possibly enjoyed the most: the "caveats" section. If some of you remember Nick presenting updates of the project at the time (e.g. BGA or WCPG between 2015 and 2017), you may remember that most of the presentation was dedicated to some of the issues discussed in the paper. For those who missed it, the "caveats" section included a sensitivity analysis of the effect of the measurement scale (IRT depression scale, raw depression sum score scale as well as a DSM-IV diagnosis in a logistic framework). Using the two depression scales we found consistent interaction results, though the strength of the interaction varied greatly. To note, the interaction did not reach significance (p=0.059) when using the DSM diagnosis, though the sample size (hence power) was lower. The analyses stratified by sex did not return significant although the statistical power was also lower. We also performed a Jinks & Fulker analysis on the MZ pairs (Jinks and Fulker 1970), with results consistent with the presence of interaction (and scale effect). Further checks included investigating the source of the interaction by separating "passive" from "active" life events (Plomin et al. 1990), or by acknowledging that the stressful life events have a genetic component (Colodro-Conde et al. 2017; Kendler and Baker 2007), which prevents from directly concluding the interaction is of the GxE type (vs. GxG). This important last point was raised by the reviewers and got solved that next Tuesday during our meeting (credit goes to Sarah Medland). At 6 hands, with Sarah, it took us less than 30 mins to implement the analysis, which was all she had before her next meeting. The solution came from taking advantage of the twin sample, which was only a complication thus far, forcing us to use mixed models to account for the sample relatedness in the analyses. We fitted a multivariate twin model on the items of the stress score, which allowed partitioning the stressful life events score into a genetic and an environmental factor score. We confirmed that most of the observed interaction could be attributed to a GxE effect as opposed to a GxG effect (Colodro-Conde et al. 2017). To be exhaustive, there is one caveat we did not implement correctly, which related to controlling for all first order interactions (Keller 2014). This was pointed out by Matt Keller after the publication and we went back to the data to make sure it did not change the results.

When the paper was accepted in *Molecular Psychiatry* we all happily celebrated in Kafenio, one of Nick's favourite restaurants in Brisbane where beautiful characters serve delightful authentic Greek/Cypriot cuisine. The sense of accomplishment may have interacted with the

buzzing effect of the wine, but we will seek replication to conclude about what caused the hangover.

Nick's energy and passion in research are contagious and have inspired us to work in research and human genetics in particular. We feel deeply grateful for having witnessed it from the front row, and possibly having fuelled it at times. This experience was extremely formative and we both feel we have gained a lot more than a good publication. The epilogue of the story could be a second project we embarked on almost immediately after, which focused one the genetic relationship between schizophrenia and population density of where people live (idea of Marcella Rietschel) (Colodro-Conde et al. 2018). It included a side GxE analysis of population density, with age as a modifier, which suggested genetic control over living environment increases with age. Nick did not find this result surprising, he had already published this result (Whitfield et al. 2005).

The overall contribution of GxE to most traits is still unknown yet heavily discussed. Evidence of specific GxE interactions has been found for depression but they individually do not explain a large fraction of the depression risk (Colodro-Conde et al. 2017; Musliner et al. 2015; Peyrot et al. 2014; Arnau-Soler et al. 2019). More research is needed, so what is your next idea Nick?

References

Arnau-Soler, Aleix, Mark J. Adams, Toni-Kim Clarke, Donald J. MacIntyre, Keith Milburn, Lauren Navrady, Generation Scotland, et al. 2019. "A Validation of the Diathesis-Stress Model for Depression in Generation Scotland." *Translational Psychiatry* 9 (1): 25. https://doi.org/10.1038/s41398-018-0356-7.

Bates, Timothy C., Narelle K. Hansell, Nicholas G. Martin, and Margaret J. Wright. 2016. "When Does Socioeconomic Status (SES) Moderate the Heritability of IQ? No Evidence for G×SES Interaction for IQ in a Representative Sample of 1176 Australian Adolescent Twin Pairs." *Intelligence* 56 (May): 10–15. https://doi.org/10.1016/j.intell.2016.02.003.

Boomsma, D. I., and N. G. Martin. 2002. "Martin NG, Gene-Environment Interactions." *Biological Psychiatry*, 181–87.

"Cancer in Australia: An Overview 2012, Table of Contents - Australian Institute of Health and Welfare." n.d. Accessed February 19, 2020. https://www.aihw.gov.au/reports/cancer/cancer-in-australia-an-overview-2012/contents/table-of-contents.

Colodro-Conde, L., B. Couvy-Duchesne, John B. Whitfield, Fabian Streit, Scott Gordon, Kathryn E. Kemper, Loic Yengo, et al. 2018. "Association Between Population Density and Genetic Risk for Schizophrenia." *JAMA Psychiatry* 75 (9): 901–10. https://doi.org/10.1001/jamapsychiatry.2018.1581.

Colodro-Conde, L., B. Couvy-Duchesne, G. Zhu, W. L. Coventry, E. M. Byrne, S. Gordon, M. J. Wright, et al. 2017. "A Direct Test of the Diathesis-Stress Model for Depression." *Molecular Psychiatry* 23 (7): 1590–96. https://doi.org/10.1038/mp.2017.130.

Cornes, B. K., P. A. Lind, S. E. Medland, G. W. Montgomery, D. R. Nyholt, and N. G. Martin. 2009. "Replication of the Association of Common Rs9939609 Variant of FTO with Increased BMI in an Australian Adult Twin Population but No Evidence for Gene by Environment (G x E) Interaction." *International Journal of Obesity (2005)* 33 (1): 75–79. https://doi.org/10.1038/ijo.2008.223.

D'Haenen, Hugo A. H., ed. 2002. *Biological Psychiatry*. Chichester: Wiley.

Dunn, Erin C., Anna Wiste, Farid Radmanesh, Lynn M. Almli, Stephanie M. Gogarten, Tamar

Sofer, Jessica D. Faul, et al. 2016. "Genome-Wide Association Study (GWAS) and Genome-Wide Environment Interaction Study (GWEIS) of Depressive Symptoms in African American and Hispanic/Latina Women." *Depression and Anxiety* 33 (4): 265–80. https://doi.org/10.1002/da.22484.

Eaves, L. J., Krystyna Last, N. G. Martin, and J. L. Jinks. 1977. "A Progressive Approach to Non-Additivity and Genotype-Environmental Covariance in the Analysis of Human Differences." *British Journal of Mathematical and Statistical Psychology* 30 (1): 1–42. https://doi.org/10.1111/j.2044-8317.1977.tb00722.x.

Gillespie, N., K. M. Kirk, A. C. Heath, N. G. Martin, and I. Hickie. 1999. "Somatic Distress as a Distinct Psychological Dimension." *Social Psychiatry and Psychiatric Epidemiology* 34 (9): 451–58. https://doi.org/10.1007/s001270050219.

Heath, Andrew C., Lindon J. Eaves, and Nicholas G. Martin. 1998. "Interaction of Marital Status and Genetic Risk for Symptoms of Depression." *Twin Research and Human Genetics* 1 (3): 119–22. https://doi.org/10.1375/twin.1.3.119.

Heath, Andrew C., Alexandre A. Todorov, Elliot C. Nelson, Pamela A. F. Madden, Kathleen K. Bucholz, and Nicholas G. Martin. 2002. "Gene-Environment Interaction Effects on Behavioral Variation and Risk of Complex Disorders: The Example of Alcoholism and Other Psychiatric Disorders." *Twin Research: The Official Journal of the International Society for Twin Studies* 5 (1): 30–37. https://doi.org/10.1375/1369052022875.

Jinks, J. L., and D. W. Fulker. 1970. "Comparison of the Biometrical Genetical, MAVA, and Classical Approaches to the Analysis of Human Behavior." *Psychological Bulletin* 73 (5): 311–49. https://doi.org/10.1037/h0029135.

Keller, Matthew C. 2014. "Gene × Environment Interaction Studies Have Not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution." *Biological Psychiatry* 75 (1): 18–24. https://doi.org/10.1016/j.biopsych.2013.09.006.

Kendler, Kenneth S., and Jessica H. Baker. 2007. "Genetic Influences on Measures of the Environment: A Systematic Review." *Psychological Medicine* 37 (5): 615–26. https://doi.org/10.1017/S0033291706009524.

Kirk, K. M., A. J. Birley, D. J. Statham, B. Haddon, R. I. Lake, J. G. Andrews, and N. G. Martin. 2000. "Anxiety and Depression in Twin and Sib Pairs Extremely Discordant and Concordant for Neuroticism: Prodromus to a Linkage Study." *Twin Research: The Official Journal of the International Society for Twin Studies* 3 (4): 299–309. https://doi.org/10.1375/136905200320565274.

Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Stephan Ripke, Naomi R. Wray, Cathryn M. Lewis, Steven P. Hamilton, Myrna M. Weissman, Gerome Breen, et al. 2013. "A Mega-Analysis of Genome-Wide Association Studies for Major Depressive Disorder." *Molecular Psychiatry* 18 (4): 497–511. https://doi.org/10.1038/mp.2012.21.

Martin, N. G. 1976. "Aspects of Human Behaviour Genetics." England: Department of Genetics, University of Birmingham.

Martin, N. G., L. J. Eaves, and A. C. Heath. 1987. "Prospects for Detecting Genotype X Environment Interactions in Twins with Breast Cancer." *Acta Geneticae Medicae Et Gemellologiae* 36 (1): 5–20. https://doi.org/10.1017/s0001566000004542.

Musliner, K. L., F. Seifuddin, J. A. Judy, M. Pirooznia, F. S. Goes, and P. P. Zandi. 2015. "Polygenic Risk, Stressful Life Events and Depressive Symptoms in Older Adults: A Polygenic Score Analysis." *Psychological Medicine* 45 (8): 1709–20. https://doi.org/10.1017/S0033291714002839.

Peyrot, Wouter J., Yuri Milaneschi, Abdel Abdellaoui, Patrick F. Sullivan, Jouke J. Hottenga, Dorret I. Boomsma, and Brenda W. J. H. Penninx. 2014. "Effect of Polygenic Risk Scores on Depression in Childhood Trauma." *The British Journal of Psychiatry: The*

Journal of Mental Science 205 (2): 113–19. https://doi.org/10.1192/bjp.bp.113.143081.

Plomin, R., P. Lichtenstein, N. L. Pedersen, G. E. McClearn, and J. R. Nesselroade. 1990. "Genetic Influence on Life Events during the Last Half of the Life Span." *Psychology and Aging* 5 (1): 25–30. https://doi.org/10.1037//0882-7974.5.1.25.

Purcell, Shaun. 2002. "Variance Components Models for Gene-Environment Interaction in Twin Analysis." *Twin Research: The Official Journal of the International Society for Twin Studies* 5 (6): 554–71. https://doi.org/10.1375/136905202762342026.

Siskind, Victor, Joanne Aitken, Adèle Green, and Nicholas Martin. 2002. "Sun Exposure and Interaction with Family History in Risk of Melanoma, Queensland, Australia." *International Journal of Cancer* 97 (1): 90–95. https://doi.org/10.1002/ijc.1563.

Slutske, Wendy S., Arielle R. Deutsch, Dixie J. Statham, and Nicholas G. Martin. 2015. "Local Area Disadvantage and Gambling Involvement and Disorder: Evidence for Gene-Environment Correlation and Interaction." *Journal of Abnormal Psychology* 124 (3): 606–22. https://doi.org/10.1037/abn0000071.

Staples, Margaret P., Mark Elwood, Robert C. Burton, Jodie L. Williams, Robin Marks, and Graham G. Giles. 2006. "Non-Melanoma Skin Cancer in Australia: The 2002 National Survey and Trends since 1985." *The Medical Journal of Australia* 184 (1): 6–10.

Treloar, S. A., N. G. Martin, K. K. Bucholz, P. A. Madden, and A. C. Heath. 1999. "Genetic Influences on Post-Natal Depressive Symptoms: Findings from an Australian Twin Sample." *Psychological Medicine* 29 (3): 645–54. https://doi.org/10.1017/s0033291799008387.

Whitfield, John B., Gu Zhu, Andrew C. Heath, and Nicholas G. Martin. 2005. "Choice of Residential Location: Chance, Family Influences, or Genes?" *Twin Research and Human Genetics* 8 (1): 22–26. https://doi.org/10.1375/twin.8.1.22.

Wray, Naomi R., Michael E. Goddard, and Peter M. Visscher. 2007. "Prediction of Individual Genetic Risk to Disease from Genome-Wide Association Studies." *Genome Research* 17 (10): 1520–28. https://doi.org/10.1101/gr.6665407.

Yang, Jian, Noah A. Zaitlen, Michael E. Goddard, Peter M. Visscher, and Alkes L. Price. 2014. "Advantages and Pitfalls in the Application of Mixed Model Association Methods." *Nature Genetics* 46 (2): 100–106. https://doi.org/10.1038/ng.2876.

Cynthia Bulik

Nick was a force of nature when it came to recruiting participants for the Anorexia Nervosa Genetics Initiative (ANGI) in Australia. This study, funded by the Klarman Family Foundation, formed the backbone of the *Nature Genetics* paper **Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. PMID:31308545.** Even though eating disorders is not his primary area of expertise, Nick managed to garner support from clinicians, families, and individuals with the illness around Australia to build a community of participation that made the study possible. It is not always easy as an outsider to step into our field and have an immediate positive impact, but Nick did. His success with ANGI is the main reason why Australia is now a site for our next major endeavor, the Eating Disorders Genetics Initiative (EDGI). I imagine he will have the same impact again! Nick is my role model for collegiality, productivity, and generativity in the latter stages of one's career.

Perhaps even more impressive, is a trait that I knew Nick had within his field, but I was astonished to see that it generalized outside of behavior genetics. Nick travels a lot, so he is often circadian-ly challenged. This means that he may occasionally doze off during talks. I had witnessed his uncanny ability to sleep through a talk only to wake up during the Q&A period to ask the single most pertinent question—almost as if he were sleeping with one ear open. I wrote this off to just being a giant in the BG field, until I saw him do it at an eating disorders conference on a topic that was completely outside of his area of expertise! The only other person I have ever seen be able to pull this off is Lindon Eaves and I am convinced it is a sign of genius. Or it is a total con and he just wants us to think that he is sleeping!

Martin Kennedy

I'd known Nick as a force in the field of human genetics, but not directly worked with him, until the ANGI study came along. I recall Cindy emailed me one day to ask if anyone was doing eating disorders research in New Zealand, and I suspect this may have been partly prompted by Nick. I'm a geneticist and knew little about eating disorders, but had a very good colleague in Jenny Jordan, a clinical psychologist, who did. Jenny and I wrote and won a small research grant that funded us to get going and contribute samples to the Australian arm of the study, and of course this brought us into close proximity with Nick and his team. Once Nick realised we were committed to this, he created the opportunity for us to continue contributing samples and provided resources for us to achieve that. It was a steep learning curve for us, but Nick's generosity, wisdom and enthusiasm ensured we worked hard to meet the study goals. Of course, this is something so many will have experienced, and must surely be part of the secret of Nick's enormous success. He is adept at charming, motivating, exciting, and if need be, pressurising people to perform. He is wonderfully collegial, witty and has a highly tuned ability to rapidly sort the wheat from the chaff. Combined with his expansive knowledge and prodigious work output, this makes for an extraordinary package.

Nick called me a year or two ago in conjunction with a depression GWAS he was running, and which he'd hoped we would be able to contribute to (sadly we were unable to). He said to me then that he got great pleasure from organising and leading large GWAS consortia, realised it was something he was very good at, and felt this might be something of a lasting legacy from the later years of his career.

I should add it is never dull being at a conference when Nick is present. At question time he can be provocative or challenging, but always with good humour and sincerity. He is also generous with praise and support, particularly for young investigators. More than anything, his love of science, and of discussing science, shines through. He is an extraordinary role model for all of us!

We are very proud that New Zealand is part of the EDGI Team, but I do not think we would have been able to take this on without the ongoing support, expertise and resources that Nick and his QIMR Berghofer team have provided. I suspect this is probably the case for many smaller groups contributing to large GWAS studies in this part of the world.

Tracey Wade
I first met Nick on the banks of the Torrens River outside of the University of Adelaide, almost 30 years ago, to discuss his involvement as a supervisor of my PhD, to be conducted at Flinders University. I had the idea that I would like to investigate the role of genes in eating disorders, as there was very little work in this area. As a clinical psychologist I really had no idea what I was talking about, but Nick was immediately enthusiastic, even more so when he found out that my husband was a member of the Liberal Party. Despite his best efforts to induce me to relocate to Brisbane, I stayed in Adelaide for the duration of the thesis. Looking back now as a seasoned PhD supervisor, I marvel anew at Nick's generosity and support over that time. He has been one of a handful of extremely influential supervisors in my life (of whom Cindy Bulik was another), who challenged me to aim higher and do better and answer questions with scientific rigour, a legacy that has stood me well over my career, and continues to inspire me in the mentoring I give to my own research students.

Nick's mentorship did not stop with my PhD. He actively pursued a postdoctoral opportunity for me with Kenneth Kendler and Cynthia Bulik at the Virginia Institute for Psychiatric and Behavioral Genetics of Virginia Commonwealth University. In recent years he supported my nomination to the Academy of Social Sciences in Australia. Over my career, we have published many papers together, mainly in the behavioural genetics of eating disorders, culminating in the *Nature Genetics* ANGI paper last year.

The work that Nick has been involved in has made a major contribution to moving the field from blaming the family for causing an eating disorder, to starting to appreciate the complexity of the various genetic and (mainly) non-shared environmental risk factors. This has had a critical impact on the field, allowing us to devise non-blaming therapies and work with families as partners in treatment.

Nick is an impressive poster boy for valuing the progress of science above all else and paying no heed to the territorial and ownership issues which can plague some scientific endeavour. He was Open Science before it was even a thing. Interactions with him are always energising (even, as Cindy points out, when he is ostensibly sleeping). Those of us who are rather less

brilliant and extroverted can feel rather drained after some encounters, but they always make our work richer, and inspire us to do better.

(Nick, Cynthia Bulik and June Alexander)

Martin HC, Christ R, Hussin JG, O'Connell J, Gordon S, Mbarek H, Hottenga JJ, McAloney K, Willemsen G, Gasparini P, Pirastu N, Montgomery GW, Navarro P, Soranzo N, Toniolo D, Vitart V, Wilson JF, Marchini J, Boomsma DI, Martin NG, Donnelly P (2015). Multicohort analysis of the maternal age effect on recombination. *Nat Commun*, 6, 7846. (cites = 15)

Niemi MEK, Martin HC, Rice DL, Gallone G, Gordon S, Kelemen M, McAloney K, McRae J, Radford EJ, Yu S, Gecz J, Martin NG, Wright CF, Fitzpatrick DR, Firth HV, Hurles ME, Barrett JC (2018). Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature*, 562(7726), 268-271. (cites = 53)

Nature via Nurture

Hilary Martin

I have been asked to write a piece about my father's role as a scientific mentor. This is impossible for me to dissociate from his role as my dad, but since neither of us would like it if I wore my heart on my sleeve, I'll be treading a fine line.

Perhaps incongruously for someone who is such a firm believer in the relative importance of genetics over environment for positive life outcomes, my father was devoted to nurturing his daughters in all things scientific, cultural, culinary and disciplinary. Maths homework was as important as piano practice or learning how to roast a leg of lamb. My earliest genetics lessons with him started over the dinner table probably around the age of 10, when he would explain the twin method by drawing structural equation models on the backs of shopping lists. I don't think I fully understood these at the time (and indeed, still don't today), but it wasn't hard to be infected by his huge enthusiasm for the subject and his conviction that being a geneticist was the most exciting job one could have. I don't think he had ever seriously contemplated any alternative career paths, and although I briefly flirted with the idea of becoming an historian, by the age of 15, I was pretty convinced that I, too, would be a geneticist, despite having taken no formal genetics classes yet.

On his firm advice, I steered clear of high school 'muddy pond' biology in favour of Latin, which he considered essential for teaching one logic and grammar, and understanding Western history, art and culture (he was right). I was introduced to the terror of his famous red pen over my Latin and history essays. Never a single ambiguity, tautology or trendy phrase slipped by him without comment, and I quickly had to learn his unique shorthand notes. I was always relieved to see the annotation "*stat*" [Latin for "it stands"], meaning "ignore this crossed-out phrase, actually it's fine as is". Similarly, his regular email correspondence will be familiar with his tendency to use brief Latin, French or German words whenever it saves having to type a few more letters painstakingly with two fingers.

In my earliest undergraduate genetics lectures in 2007, his prophecy came true, and I discovered that biology only became interesting when one introduced some maths. I felt the same excitement that he had experienced at the same stage upon understanding how a binomial distribution could converge to a normal distribution, bridging Mendelian and complex genetics. I remember him coming home one evening that year, throwing the first WTCCC GWAS paper down on the table, and declaring 'this is a really important paper'. It was a thrilling time to be starting out in the field. Like many Australian undergraduates, I lived at home which meant I didn't have the same riotous fun and new social experiences as other students. However, the frequent boozy dinner parties with my parents and their colleagues and the many conversations about the latest exciting developments in the field were formative, and a lot of fun. These conversations often degenerated into political debates in which my father would make more and more outrageous arguments to wind people up, as the wine flowed freely. Our less practiced guests would become duly enraged by his extreme positions, but the seasoned old-timers like Matt Keller, Manuel Ferreira, Naomi Wray, Peter Visscher, and Nathan Gillespie, would just shake their heads and laugh, knowing that he didn't really mean what he was saying (well, at least not 100%). My father would take great pleasure in sharing the latest fun facts from his work, such as how the first hair curliness gene they found in humans (trichohyalin) was also involved in wool crimping in sheep, or how UK Biobank data have demonstrated a negative genetic correlation between playing computer games and IQ but also, ironically, with the time to click through the IQ test.

We have a few papers together to date, one on patterns of recombination in human pedigrees (PMID 26242864) and another on the contribution of common variants to rare developmental disorders (PMID 30258228). On both occasions, my father fell over backwards to help in providing the data, only being kept from breaking too many rules by his dedicated staff, to whom he is eternally loyal. He has never had any patience for, as he sees it, the sanctimonious proponents of 'open data sharing', pointing out, correctly, that he has always shared his data with any even vaguely competent researcher who asked, but on his own terms, under his guidance, so he can ensure people understand the ins-and-outs of the

data collection and how variables can be missing-not-at-random etc. My mother and I roll our eyes when we watch him compulsively working the room at conferences, starting five new collaborations an hour – we know it will lead to more complaints about how all he does is send emails and red-pen manuscripts, although he does love it, really.

I have attended the Boulder course twice now to help with teaching, and it has been a joy to see my father in his element there, with some of his protégés now running the workshop. He is never afraid to ask a naïve question in front of the whole room to make a didactic point, which I think many students (and faculty members) appreciate. One of the most important things he has taught me is that there is no such thing as a stupid question, and that you will always learn more by putting yourself out there, asking it and digesting the answer, than by sitting back passively and never quite understanding. If I'm half as active and passionate about the subject at his age, it will be a great testament to his influence.

(Christmas carols with glass of wine in hand. Hilary Martin with Nick plus Beryl and Felicity Martin)