

## Genome-wide association studies and the genetics of entrepreneurship

Matthijs J. H. M. van der Loos ·  
Philipp D. Koellinger · Patrick J. F. Groenen ·  
A. Roy Thurik

Received: 9 December 2009 / Accepted: 15 December 2009 / Published online: 8 January 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

Entrepreneurship is a vital element of well-functioning economies. It is sometimes denoted as their ‘scarcest input factor’. Entrepreneurs introduce innovations into the economic system and may contribute towards higher productivity levels and hence economic growth [1, 2, 3]. In addition, market entry by entrepreneurial activity is vital in adjusting markets towards competitive levels [4], and even purely imitative entrepreneurial activity can have growth-enhancing effects by stimulating efficiency and promoting the diffusion of technologies [5]. Hence, understanding the reasons why and under which circumstances people engage in entrepreneurial activity is important. Traditionally, research on the determinants of entrepreneurship has focused on factors that are easy to observe, such as socio-demographics. Different preferences of people are also used to explain the interpersonal variation in entrepreneurial activity. More recently, research found that entrepreneurs often exhibit different cognitive processes that result in different perceptions and interpretations of themselves and their environment [6, 7]. While economics helps us understand the complex interactions between individuals and environmental conditions that ultimately result in behaviour, the relevance of individual differences in preferences, cognition, and personality raises the question if

genetic variation could be relevant in explaining economic decisions. Indeed, a recently published twin study suggests that genetic differences among people can influence their tendency to become entrepreneurs [8]. The potential relevance of genes in economic behaviour raises various new research questions, including which interactions of genes and environmental conditions tend to result in particular outcomes; how people with particular genes fit with given environments or self-select into them; and how the interplay of individuals and their environment results in prosperity and satisfaction of people or a lack thereof.

Fuelled by technological developments from the Human Genome and HapMap projects, an unprecedented era of genetic discoveries has been launched by the application of the genome-wide association (GWA) design. Genome-wide association studies (GWAS) (now with >400 published studies) have been successful in identifying common variants associated with numerous complex quantitative traits and diseases [9]. GWAS focus on single nucleotide polymorphisms (SNPs) covering a high proportion of the common genetic variation in the genome.

The first GWAS used only 10,000 genotyped SNPs in 100 individuals [10], but the field has evolved enormously. Decreasing genotyping costs and improved statistical techniques have made it possible to analyse up to 1 million genotyped and 2.5 million imputed SNPs. In the near future, it is expected that the number of different SNPs that can be genotyped will be 2–12 million. However, with the increase in the number of SNPs and consequently the number of statistical tests it can be expected on the basis of pure chance that a large number of SNPs will show significant associations. For example, assume that none of the analysed 500,000 SNPs is associated with an outcome, i.e., that the statistical null hypothesis is correct. If we adopt a 1% significance level for hypothesis testing, performing

---

M. J. H. M. van der Loos (✉) · P. D. Koellinger · A. R. Thurik  
Department of Applied Economics, Erasmus School  
of Economics, Erasmus University Rotterdam,  
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands  
e-mail: mvanderloos@ese.eur.nl

P. J. F. Groenen  
Econometric Institute, Erasmus School of Economics,  
Erasmus University Rotterdam, P.O. Box 1738,  
3000 DR Rotterdam, The Netherlands

500,000 tests will yield 5,000 expected incorrect rejections of the null hypothesis. Hence, to keep the false positive rate at an acceptable level, very stringent significance levels are required in GWAS to adjust for multiple testing. The often used Bonferroni correction, for example, suggests a  $P$  value of smaller than  $2 \times 10^{-8}$  if the significance level for the whole family of 500,000 tests is supposed to be 1%. To be able to discover associations with weak effects, very large sample sizes are needed [11]. As a consequence, collaborative research consortia have been assembled to share GWAS data usually analysed in the form of meta-analysis. The large sample sizes and replication of associations therein most likely reflect that genome-wide significant findings are true positives.

We assembled a multidisciplinary research group of economists and (genetic) epidemiologists focused on testing whether relatively general economic behaviours—like becoming an entrepreneur—can be influenced by genes. To the best of our knowledge, this is the earliest attempt to apply GWAS to an economic outcome of a relatively general nature and will reveal potentials and limitations of this approach for economic research. There is also potential of our research approach to inform medical research: since becoming an entrepreneur or not affects income [12, 13], life-style [14, 15], and happiness [16, 17, 18], this choice can in turn influence medical conditions. In general, a mismatch between the genetic predisposition of people and their actual working conditions could result in unfavourable health, depending on the genetic ‘fit’ between individuals and their working conditions. In addition, a lack of desired social status seems to be associated with earlier death [19].

The first challenge was to define an accurate phenotype definition. As entrepreneurship is a phenomenon that can materialise in many different forms, different definitions and operationalisations coexist [20]. We have opted to operationalise entrepreneurship as self-employment within the setting of the Rotterdam Study [21]. The Rotterdam Study is a prospective cohort study, hosted at the Erasmus Medical Center, and started with a pilot phase in the second half of 1989. From January 1990 to September 1993, 7,983 participants were successfully recruited in the well-defined Ommoord district in Rotterdam. This formed the initial cohort called Rotterdam Study I (RS-I). The participants were all 55 years of age or over when entering the study and the oldest participant at the start was 106 years. From February 2000 until December 2001, an additional 3,011 participants were interviewed and gathered within a second cohort: Rotterdam Study II (RS-II). The participants consisted of individuals who became 55 years since the initial study or those of 55 years and older who moved into the Ommoord district. The study was again extended from February 2006 until December 2008 with a third cohort, Rotterdam Study III (RS-III), consisting of 3,932

individuals of 45 years and older living in the district and who had not been previously interviewed. This last extension increased the number of participants of the Rotterdam Study to a total of 14,926. The majority of the genotyped individuals in the Rotterdam Study provided data on their complete working life histories *and* whether they were self-employed during any of their occupations. An explicit advantage of using a sample of elderly individuals is that most uncertainties about future occupations of the respondents are resolved since a large part of the sample had already reached the official retirement age in the Netherlands of 65 years, which allows us to look back at the work life histories of these people. Based on this information, we can differentiate between respondents who were never self-employed (control group), at least once self-employed, serial self-employed, and never anything else than self-employed. Thus, we can differentiate in the discovery sample between different degrees of entrepreneurial activity.

We presented preliminary findings from our discovery cohort at the Behavior Genetics Association in Louisville, Kentucky, in June 2008. Our work since then focuses on replicating results in independent samples and we have now embedded our effort to assemble a working group within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium [22]. The CHARGE consortium consists of the following five independent cohort studies: the Age, Gene, Environment, Susceptibility Study (AGES), the Atherosclerosis Risk in Communities Study (ARIC), the Cardiovascular Health Study (CHS), the Framingham Heart Study (FHS), and the Rotterdam Study (RS). Together these studies provide follow-up data on 50,000–70,000 individuals from US and European ancestry. The ongoing plan is to recruit additional cohorts with data on entrepreneurship and extend our discovery sample to achieve a sufficiently-powered setting to identify common genetic variants underlying the propensity to become an entrepreneur. To this end we are setting up a consortium we have termed the ‘Entrepreneur Consortium’, which already includes the St Thomas’ UK Adult Twin Registry [23] and the Netherlands Twin Register [24] and will include the aforementioned CHARGE cohorts. Additionally, a collaboration with the Erasmus Rucphen Family study (ERF) [25] is being set up. An extended description of the study setup is forthcoming [26]. Finally, our consortium also aims to set the well-powered stage to perform more extensive genetic- as well as biologically-oriented studies into entrepreneurship.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Carree M, Thurik R. *Entrepreneurship and economic growth*. Cheltenham: Edward Elgar; 2006.
2. Audretsch D, Keilbach M. Entrepreneurship capital and economic performance. *Reg Stud*. 2004;38:949–59.
3. van Stel A, Carree M, Thurik R. The effect of entrepreneurial activity on national economic growth. *Small Bus Econ*. 2005;24:311–21.
4. Kirzner IM. *Competition and entrepreneurship*. Chicago and London: The University of Chicago Press; 1973.
5. Schmitz JA Jr. Imitation, entrepreneurship, and long-run growth. *J Polit Econ*. 1989;97:721–39.
6. Camerer C, Lovallo D. Overconfidence and excess entry: an experimental approach. *Am Econ Rev*. 1999;89:306–18.
7. Koellinger P, Minniti M, Schade C. “I think I can, I think I can”: overconfidence and entrepreneurial behavior. *J Econ Psychol*. 2007;28:502–27.
8. Nicolaou N, Shane S, Hunkin J, Cherkas L, Spector T. Is the tendency to engage in entrepreneurship genetic? *Manag Sci*. 2008;54:167–79.
9. Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA*. 2009;106:9362–67.
10. Hu N, Wang C, Hu Y, Yang HH, Giffen C, Tang ZZ, et al. Genome-wide association study in esophageal cancer using GeneChip mapping 10K array. *Cancer Res*. 2005;65:2542–6.
11. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9:356–69.
12. Hamilton BH. Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *J Polit Econ*. 2000;108:604–31.
13. Moskowitz TJ, Vissing-Jørgensen A. The returns to entrepreneurial investment: a private equity premium puzzle? *Am Econ Rev*. 2002;92:745–78.
14. Hyttinen A, Ruuskanen OP. Time use of the self-employed. *Kyklos*. 2009;60:105–22.
15. Parasuraman S, Simmers CA. Type of employment, work-family conflict and well-being: a comparative study. *J Organ Behav*. 2001;22:551–68.
16. Benz M, Frey BS. Being independent is a great thing: subjective evaluations of self-employment and hierarchy. *Economica*. 2008;75:362–83.
17. Blanchflower DG, Oswald AJ. What makes an entrepreneur? *J Labor Econ*. 1998;16:26–60.
18. Block J, Koellinger P. I can’t get no satisfaction—necessity entrepreneurship and procedural utility. *Kyklos*. 2009;62:191–209.
19. Rablen MD, Oswald AJ. Mortality and immortality: the Nobel Prize as an experiment into the effect of status upon longevity. *J Health Econ*. 2008;27:1462–71.
20. Shane S, Venkataraman S. The promise of entrepreneurship as a field of research. *Acad Manage Rev*. 2000;25:217–26.
21. Hofman A, Breteler MMB, van Duijn CM, Janssen HLA, Krestin GP, Kuipers EJ, et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol*. 2009;24:553–72.
22. Psaty BM, O’Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JJ, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet*. 2009;2:73–80.
23. Spector TD, Williams FMK. The UK Adult Twin Registry (TwinsUK). *Twin Res Hum Genet*. 2006;9:899–906.
24. Boomsma DI, de Geus EJC, Vink JM, Stubbe JH, Distel MA, Hottenga JJ, et al. Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet*. 2006;9:849–57.
25. Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet*. 2005;69:288–95.
26. Koellinger PD, van der Loos MJHM, Groenen PJF, Hofman A, Rivadeneira F, Thurik AR, et al. Genome-wide association studies in economic and entrepreneurship research: promises and limitations. *Small Bus Econ*. 2010 (Forthcoming).