THE UNIVERSITY OF CHICAGO

PARENT OF ORIGIN EFFECTS ON GENE EXPRESSION AND QUANTITATIVE TRAITS IN THE HUTTERITES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON GENETICS

BY
SAHAR VICTORIA MOZAFFARI

CHICAGO, ILLINOIS
2018

Table of Contents

List of Figures

List of Tables [1]

1. Note: Due to the large size of some tables, the tables have been provided in a supplementary file accompanying the dissertation. In such cases, the page number provided below directs the reader to a table's caption.

ACKNOWLEDGMENTS

ABSTRACT

CHAPTER 1

INTRODUCTION

1.1 Human Genetics and the Genetics of Complex Traits

A central goal of genetics is to understand the contribution of genetic variation to phenotypic variation.

1.2 The Origin of Genomic Imprinting

Genomic imprinting in its broadest sense suggests that a phenotype observed for a particular gene or genes depends on the sex of the parent from with the gamete containing that gene or genes originated[22]. It was said that a particular gene is imprinted if it results in a different phenotype when it is maternally inherited versus paternally inherited.

The first use of the term "imprinting" was used in reference to the recognition by the cell of of chromosomes in *Sciara*[7, 22]. "The "imprint" a chromosome bears is unrelated to the genic constitution of the chromosome and is determined only by the sex of the germ line through with the chromosome has been inherited."[7]

The preferential inactivation of the paternally-derived X chromosomes in mouse were the first demonstrations of a functional imprint in mammalian genomes[23, 17, 5]. The first suggestion of imprinting on autosomes was by a deletion on mouse chromosome 17 that showed a different phenotype based on which parent the deletion was inherited from.[14, 13].

It was not until the development of the pronuclear transplantation technique that allowed for the creation of mice zygotes which contained only maternal or only paternal genetic contributions and provided evidence that the maternal and paternal genomes are not equal. The differential imprinting on the parental chromosomes prevented complete embryonic development in these mice with complete uniparental disomy.[22, 18].

Figure 1.1: **Asthma GWAS - low hanging fruit**

1.3 The Search for Parent of Origin Effects

CHAPTER 2

PARENT OF ORIGIN EFFECTS ON QUANTITATIVE PHENOTYPES IN A FOUNDER POPULATION

2.1 Abstract[1]

The impact of the parental origin of associated alleles in GWAS has been largely ignored. Yet sequence variants could affect traits differently depending on whether they are inherited from the mother or the father. To explore this possibility, we studied 21 quantitative phenotypes in a large Hutterite pedigree. We first identified variants with significant single parent (maternal-only or paternal-only) effects, and then used a novel statistical model to identify variants with opposite parental effects. Overall, we identified parent of origin effects (POEs) on 11 phenotypes, most of which are risk factors for cardiovascular disease. Many of the loci with POEs have features of imprinted regions and many of the variants with POE are associated with the expression of nearby genes. Overall, our results indicate that POEs, which can be opposite in direction, are relatively common in humans, have potentially important clinical effects, and will be missed in traditional GWAS.

2.2 Introduction

Genome-wide association studies (GWAS) typically treat alleles inherited from the mother and the father as equivalent, although variants can affect traits differently depending on whether they are maternal or paternal in origin. In particular, parent of origin effects (POEs) can result from imprinting, where epigenetic modifications allows for differential gene expression on homologous chromosomes that is determined by the parental origin of the chromosome. Mutations in imprinted genes or regions can result in diseases. For example,

1. Citation for chapter:

two very different diseases, Prader-Willi Syndrome (PWS) and Angelman Syndrome (AS), are due to loss of function alleles in genes within an imprinted region on chromosome 15q11-13. Inheriting a loss of function mutation for the SNRPN gene from the father results in PWS but inheriting a loss of function mutation for the UBE3A gene from the mother results in AS[21, 9] Long noncoding RNA genes at this and other imprinted regions act to silence (i.e. imprint) genes in cis. Imprinted genes are often part of imprinted gene networks, suggesting regulatory links between these genes [19, 10, 24]. More than 200 imprinted loci have been described in humans [4] but there are likely many other, as yet undiscovered, imprinted loci.

Previous studies have utilized pedigrees to test maternal and paternal alleles separately for association with phenotypes or with gene expression to uncover new imprinted loci [15, 3, 11, 20, 4]. Kong *et al* [15] discovered one locus associated with breast cancer risk only when the allele is inherited from the father and another locus associated with type 2 diabetes risk only when the allele is inherited from the mother. Garg et al. reported parent-of-origin cis-eQTLs with known or putative novel imprinted genes affecting gene expression[11]. Two additional studies by Zoledziewsk et al. and Benonisdottir et al. identified opposite POEs on adult height at known imprinted loci[28, 4]. Both studies reported associations with variants at the KCNQ1 gene, and one showed additional opposite POEs with height at two known imprinted loci (IGF2-H19 and DLK1-MEG3)[4]These studies provide proof-of-principle that alleles at imprinted loci can show POEs, some with opposite effects, with common phenotypes.

Many existing studies and methods identify parent of origin effects use case/parent trios or case/mother duos[6, 12, 2, 26, 25]. Similar to Kong *et al.* [15], our method does not require data on the parent and only uses the parent of origin informative alleles which were assigned and phased using PRIMAL[16]. In contrast to Kong *et al.* [15] which used binary traits, our method tests for parent of origin effects on quantitative traits, similar to Benonisdottir *et al.* [4] which tested for parent of origin effects on height.

No previous study has included a broad range of human quantitative phenotypes or has studied genome-wide variants with effects in different directions depending on the parent of origin. To address this possibility, we developed a statistical model that directly compares the effects of the maternal and paternal alleles to identify effects that are different, including those that are opposite. We applied this model in a study of 21 common quantitative traits that were measured in the Hutterites, a founder population of European descent for which we have phased genotype data [16] We identified variants with maternally inherited or paternally inherited effects only and variants with opposite POEs. Some of the identified regions have characteristics similar to known imprinted genes. Overall, we show that this model can identify putative novel imprinted regions with POEs for a broad range of clinically relevant quantitative phenotypes.

2.3 Results

2.3.1 GWAS

2.3.2 Parent of Origin GWAS

2.4 Methods

2.4.1 Sample Composition

The individuals in this study have participated in one or more of our studies on the genetics of complex traits in the Hutterites[8, 27, 1]. The more than 1,500 Hutterites in our study are related to each other in a 13-generation pedigree including 3,671 individuals.

2.4.2 Genotype Data

Variants detected in the whole genome sequences of 98 Hutterites were previously imputed to an additional 1,317 individuals using PRIMAL, a high-accuracy pedigree based imputation method[16]. PRIMAL was used to phase alleles and assign parent of origin for 83% of about 12 million autosomal SNPs. For these studies, we selected SNPs that had a MAF 1% and genotype call rate 85%. This yielded 5,891,982 autosomal SNPs. Parent of origin allele call rates differed among individuals and between phenotypes (Table S1 **??**).

2.4.3 Phenotype Data

2.4.4 GWAS

2.4.5 Maternal and Paternal GWAS

2.4.6 Differential Effect GWAS (PO-GWAS)

2.4.7 Parent of Origin eQTL studies

2.4.8 Maternal and Paternal Parent of Origin eQTL

2.4.9 Differential Parent of Origin eQTL

]

References

[1] Mark Abney, Mary Sara McPeek, and Carole Ober. Broad and Narrow Heritabilities of Quantitative Traits in a Founder Population. *The American Journal of Human Genetics*, 68(5):1302–1307, May 2001.

[2] Holly F Ainsworth, Jennifer Unwin, Deborah L Jamison, and Heather J Cordell. Investigation of maternal effects, maternal-fetal interactions and parent-of-origin effects (imprinting), using mothers and their offspring. *Genetic Epidemiology*, 35(1):19–45, December 2010.

[3] Yael Baran, Meena Subramaniam, Anne Biton, Taru Tukiainen, Emily K Tsang, Manuel A Rivas, Matti Pirinen, Maria Gutierrez-Arcelus, Kevin S Smith, Kim R Kukurba, Rui Zhang, Celeste Eng, Dara G Torgerson, Cydney Urbanek, GTEx Consortium, Jin Billy Li, Jose R Rodriguez-Santana, Esteban G Burchard, Max A Seibold, Daniel G MacArthur, Stephen B Montgomery, Noah A Zaitlen, and Tuuli Lappalainen. The landscape of genomic imprinting across diverse adult human tissues. *Genome Research*, 25(7):927–936, July 2015.

[4] Stefania Benonisdottir, Asmundur Oddsson, Agnar Helgason, Ragnar P Kristjansson, Gardar Sveinbjornsson, Arna Oskarsdottir, Gudmar Thorleifsson, Olafur B Davidsson, Gudny A Arnadottir, Gerald Sulem, Brynjar O Jensson, Hilma Holm, Kristjan F Alexandersson, Laufey Tryggvadottir, G Bragi Walters, Sigurjon A Gudjonsson, Lucas D Ward, Jon K Sigurdsson, Paul D Iordache, Michael L Frigge, Thorunn Rafnar, Augustine Kong, Gisli Masson, Hannes Helgason, Unnur Thorsteinsdottir, Daniel F Gudbjartsson, Patrick Sulem, and Kari Stefansson. Epigenetic and genetic components of height regulation. *Nature Communications*, 7:13490, November 2016.

[5] H S Chandra and S W Brown. Chromosome imprinting and the mammalian X chromosome. *Nature*, 253(5488):165–168, January 1975.

[6] Trees-Juen Chuang, Yu-Hsiang Tseng, Chia-Ying Chen, and Yi-Da Wang. Assessment of imprinting- and genetic variation-dependent monoallelic expression using reciprocal allele descendants between human family trios. *Scientific Reports*, 7(1):1–12, July 2017.

[7] H V Crouse. The Controlling Element in Sex Chromosome Behavior in Sciara. *Genetics*, 45(10):1429–1443, October 1960.

[8] Darren A Cusanovich, Minal Caliskan, Christine Billstrand, Katelyn Michelini, Claudia Chavarria, Sherryl De Leon, Amy Mitrano, Noah Lewellyn, Jack A Elias, Geoffrey L Chupp, Roberto M Lang, Sanjiv J Shah, Jeanne M Decara, Yoav Gilad, and Carole Ober. Integrated analyses of gene expression and genetic association studies in a founder population. *Human Molecular Genetics*, 25(10):2104–2112, May 2016.

[9] J Greg Falls, David J Pulford, Andrew A Wylie, and Randy L Jirtle. Genomic Imprinting: Implications for Human Disease. *The American Journal of Pathology*, 154(3):635–647, March 1999.

[10] A Gabory, M A Ripoche, A Le Digarcher, F Watrin, A Ziyyat, T Forne, H Jammes, J F X Ainscough, M A Surani, L Journot, and L Dandolo. H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. *Development*, 136(20):3413–3421, September 2009.

[11] Paras Garg, Christelle Borel, and Andrew J Sharp. Detection of Parent-of-Origin Specific Expression Quantitative Trait Loci by Cis-Association Analysis of Gene Expression in Trios. *PLoS ONE*, 7(8):e41695, August 2012.

[12] Richard Howey and Heather J Cordell. PREMIM and EMIM: tools for estimation of maternal, imprinting and interaction effects using multinomial modelling. *BMC Bioinformatics*, 13(1):149, June 2012.

[13] D R Johnson. Further observations on the hairpin-tail (Thp) mutation in the mouse. *Genetics Research*, 24(02):207–214, October 1974.

[14] D R Johnson. Hairpin-tail: a case of post-reductional gene action in the mouse egg. *Genetics*, 76(4):795–805, April 1974.

[15] Augustine Kong, Valgerdur Steinthorsdottir, Gisli Masson, Gudmar Thorleifsson, Patrick Sulem, Soren Besenbacher, Aslaug Jonasdottir, Asgeir Sigurdsson, Kari Th Kristinsson, Adalbjorg Jonasdottir, Michael L Frigge, Arnaldur Gylfason, Pall I Olason, Sigurjon A Gudjonsson, Sverrir Sverrisson, Simon N Stacey, Bardur Sigurgeirsson, Kristrun R Benediktsdottir, Helgi Sigurdsson, Thorvaldur Jonsson, Rafn Benediktsson, Jon H Olafsson, Oskar Th Johannsson, Astradur B Hreidarsson, Gunnar Sigurdsson, DIAGRAM Consortium, Anne C Ferguson-Smith, Daniel F Gudbjartsson, Unnur Thorsteinsdottir, and Kari Stefansson. Parental origin of sequence variants associated with complex diseases. *Nature*, 462(7275):868–874, December 2009.

[16] Oren E Livne, Lide Han, Gorka Alkorta-Aranburu, William Wentworth-Sheilds, Mark Abney, Carole Ober, and Dan L Nicolae. PRIMAL: Fast and Accurate Pedigree-based Imputation from Sequence Data in a Founder Population. *PLOS Computational Biology*, 11(3):e1004139, March 2015.

[17] Mary F Lyon and Sohaila Rastan. Parental source of chromosome imprinting and its relevance for X chromosome inactivation. *Differentiation*, 26(1):63–67, June 1984.

[18] James McGrath and Davor Solter. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*, 37(1):179–183, May 1984.

[19] Manus M Patten, Michael Cowley, Rebecca J Oakey, and Robert Feil. Regulatory links between imprinted genes: evolutionary predictions and consequences. *Proceedings of the Royal Society B: Biological Sciences*, 283(1824):20152760, February 2016.

[20] John Rb Perry, Felix Day, Cathy E Elks, Patrick Sulem, Deborah J Thompson, Teresa Ferreira, Chunyan He, Daniel I Chasman, Tõnu Esko, Gudmar Thorleifsson, Eva Albrecht, Wei Q Ang, Tanguy Corre, Diana L Cousminer, Bjarke Feenstra, Nora Franceschini, Andrea Ganna, Andrew D Johnson, Sanela Kjellqvist, Kathryn L Lunetta, George McMahon, Ilja M Nolte, Lavinia Paternoster, Eleonora Porcu, Albert V Smith, Lisette Stolk, Alexander Teumer, Natalia Tšernikova, Emmi Tikkanen, Sheila Ulivi, Erin K Wagner, Najaf Amin, Laura J Bierut, Enda M Byrne, Jouke-Jan Hottenga, Daniel L Koller, Massimo Mangino, Tune H Pers, Laura M Yerges-Armstrong, Jing Hua Zhao, Irene L Andrulis, Hoda Anton-Culver, Femke Atsma, Stefania Bandinelli, Matthias W Beckmann, Javier Benitez, Carl Blomqvist, Stig E Bojesen, Manjeet K Bolla, Bernardo Bonanni, Hiltrud Brauch, Hermann Brenner, Julie E Buring, Jenny Chang-Claude, Stephen Chanock, Jinhui Chen, Georgia Chenevix-Trench, J Margriet Collée, Fergus J Couch, David Couper, Andrea D Coveillo, Angela Cox, Kamila Czene, Adamo Pio D'adamo, George Davey Smith, Immaculata De Vivo, Ellen W Demerath, Joe Dennis, Peter Devilee, Aida K Dieffenbach, Alison M Dunning, Gudny Eiriksdottir, Johan G Eriksson, Peter A Fasching, Luigi Ferrucci, Dieter Flesch-Janys, Henrik Flyger, Tatiana Foroud, Lude Franke, Melissa E Garcia, Montserrat García-Closas, Frank Geller, Eco Ej de Geus, Graham G Giles, Daniel F Gudbjartsson, Vilmundur Gudnason, Pascal Guénel, Suiqun Guo, Per Hall, Ute Hamann, Robin Haring, Catharina A Hartman, Andrew C Heath, Albert Hofman, Maartje J Hooning, John L Hopper, Frank B Hu, David J Hunter, David Karasik, Douglas P Kiel, Julia A Knight, Veli-Matti Kosma, Zoltán Kutalik, Sandra Lai, Diether Lambrechts, Annika Lindblom, Reedik Mägi, Patrik K Magnusson, Arto Mannermaa, Nicholas G Martin, Gisli Masson, Patrick F McArdle, Wendy L McArdle, Mads Melbye, Kyriaki Michailidou, Evelin Mihailov, Lili Milani, Roger L Milne, Heli Nevanlinna, Patrick Neven, Ellen A Nohr, Albertine J Oldehinkel, Ben A Oostra, Aarno Palotie, Munro Peacock, Nancy L Pedersen, Paolo Peterlongo, Julian Peto, Paul Dp Pharoah, Dirkje S Postma, Anneli Pouta, Katri Pylkäs, Paolo Radice, Susan Ring, Fernando Rivadeneira, Antonietta Robino, Lynda M Rose, Anja Rudolph, Veikko Salomaa, Serena Sanna, David Schlessinger, Marjanka K Schmidt, Mellissa C Southey, Ulla Sovio, Meir J Stampfer, Doris Stöckl, Anna M Storniolo, Nicholas J Timpson, Jonathan Tyrer, Jenny A Visser, Peter Vollenweider, Henry Völzke, Gerard Waeber, Melanie Waldenberger, Henri Wallaschofski, Qin Wang, Gonneke Willemsen, Robert Winqvist, Bruce Hr Wolffenbuttel, Margaret J Wright, Australian Ovarian Cancer Study, GENICA Network, kConFab, LifeLines Cohort Study, InterAct Consortium, Early Growth Genetics (EGG) Consortium, Dorret I Boomsma, Michael J Econs, Kay-Tee Khaw, Ruth Jf Loos, Mark I McCarthy, Grant W Montgomery, John P Rice, Elizabeth A Streeten, Unnur Thorsteinsdottir, Cornelia M van Duijn, Behrooz Z Alizadeh, Sven Bergmann, Eric Boerwinkle, Heather A Boyd, Laura Crisponi, Paolo Gasparini, Christian Gieger, Tamara B Harris, Erik Ingelsson, Marjo-Riitta Jarvelin, Peter Kraft, Debbie Lawlor, Andres Metspalu, Craig E Pennell, Paul M Ridker, Harold Snieder, Thorkild Ia Sørensen, Tim D Spector, David P Strachan, André G Uitterlinden, Nicholas J Wareham, Elisabeth Widén, Marek Zygmunt, Anna Murray, Douglas F Easton, Kari Stefansson, Joanne M Murabito, and Ken K Ong. Parent-of-origin-specific

allelic associations among 106 genomic loci for age at menarche. *Nature*, 514(7520):92–97, October 2014.

[21] Jo Peters. The role of genomic imprinting in biology and disease: an expanding view. *Nature reviews. Genetics*, 15(8):517–530, August 2014.

[22] C Sapienza. Genome imprinting and dominance modification. *Annals of the New York Academy of Sciences*, 564:24–38, 1989.

[23] N Takagi and M Sasaki. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature*, 256(5519):640–642, August 1975.

[24] Annie Varrault, Charlotte Gueydan, Annie Delalbre, Anja Bellmann, Souheir Houssami, Cindy Aknin, Dany Severac, Laetitia Chotard, Malik Kahli, Anne Le Digarcher, Paul Pavlidis, and Laurent Journot. Zac1 Regulates an Imprinted Gene Network Critically Involved in the Control of Embryonic Growth. *Developmental Cell*, 11(5):711–722, November 2006.

[25] C R Weinberg, A J Wilcox, and R T Lie. A Log-Linear Approach to Case-Parent–Triad Data: Assessing Effects of Disease Genes That Act Either Directly or through Maternal Effects and That May Be Subject to Parental Imprinting. *The American Journal of Human Genetics*, 62(4):969–978, April 1998.

[26] Clarice R Weinberg. Methods for Detection of Parent-of-Origin Effects in Genetic Studies of Case-Parents Triads. *The American Journal of Human Genetics*, 65(1):229–235, July 1999.

[27] Lauren A Weiss, Mark Abney, Edwin H Cook Jr., and Carole Ober. Sex-Specific Genetic Architecture of Whole Blood Serotonin Levels. *The American Journal of Human Genetics*, 76(1):33–41, January 2005.

[28] Magdalena Zoledziewska, Carlo Sidore, Charleston W K Chiang, Serena Sanna, Antonella Mulas, Maristella Steri, Fabio Busonero, Joseph H Marcus, Michele Marongiu, Andrea Maschio, Diego Ortega Del Vecchyo, Matteo Floris, Antonella Meloni, Alessandro Delitala, Maria Pina Concas, Federico Murgia, Ginevra Biino, Simona Vaccargiu, Ramaiah Nagaraja, Kirk E Lohmueller, UK10K Consortium, Nicholas J Timpson, Nicole Soranzo, Ioanna Tachmazidou, George Dedoussis, Eleftheria Zeggini, Understanding Society Scientific Group, Sergio Uzzau, Chris Jones, Robert Lyons, Andrea Angius, Goncalo R Abecasis, John Novembre, David Schlessinger, and Francesco Cucca. Height-reducing variants and selection for short stature in Sardinia. *Nature genetics*, 47(11):1352–1356, November 2015.