

HGGA, Volume 2

Supplemental information

**Genome-wide survey of parent-of-origin-specific
associations across clinical traits derived
from electronic health records**

Hye In Kim, Bin Ye, Jeffrey Staples, Anthony Marcketta, Chuan Gao, Regeneron Genetics Center, Geisinger Regeneron DiscovEHR Collaboration, Alan R. Shuldiner, and Cristopher V. Van Hout

Supplemental Information

Geisinger Regeneron DiscovEHR Collaboration Banner and Contribution Statement

All authors/contributors are listed in alphabetical order.

Lance J. Adams¹, Jackie Blank¹, Dale Bodian¹, Derek Boris¹, Adam Buchanan¹, David J. Carey¹, Ryan D. Colonie¹, F. Daniel Davis¹, Dustin N. Hartzel¹, Melissa Kelly¹, H. Lester Kirchner¹, Joseph B. Leader¹, David H. Ledbetter¹, Ph.D., J. Neil Manus¹, Christa L. Martin¹, Michelle Meyer¹, Tooraj Mirshahi¹, Matthew Oetjens¹, Thomas Nate Person¹, Christopher Still¹, Natasha Strande¹, Amy Sturm¹, Jen Wagner¹, Marc Williams¹

Contribution: Development and validation of clinical phenotypes used to identify study participants and (when applicable) controls.

Affiliations:

1. Geisinger, Danville, PA

Regeneron Genetics Center Banner and Contribution Statements

All contributors are listed in alphabetical order. RGC Management and Leadership Team:

Goncalo R. Abecasis, D.Phil.¹, Aris Baras, M.D.¹, Michael Cantor, M.D.¹, Giovanni Coppola, M.D.¹, Aris Economides, Ph.D.¹, John D. Overton, Ph.D.¹, Jeffrey G. Reid, Ph.D.¹, Alan R. Shuldiner, M.D.¹

Contribution: All authors contributed to securing funding, study design and oversight, and review and interpretation of data and results.

Sequencing and Lab Operations:

Christina Beechert¹, Erin Brian¹, Alex DeVito¹, Caitlin Forsythe¹, Erin D. Fuller¹, Zhenhua Gu¹, Joe LaRosa¹, Michael Lattari¹, Alexander Lopez¹, Kia Manoochehri¹, Justin Marcovici¹, Manasi Pradhan¹, John D. Overton, Ph.D.¹, Thomas D. Schleicher¹, Maria Sotiropoulos Padilla¹, Karina Toledo¹, Emelia Weihinig¹, Louis Widom¹, Sarah E. Wolf¹, Ricardo H. Ulloa¹

Contribution: Performed and are responsible for sample genotyping and exome sequencing, conceived and are responsible for laboratory automation, and responsible for sample tracking and the library information management system.

Genome Informatics:

Xiaodong Bai, Ph.D.¹, Suganthi Balasubramanian, Ph.D.¹, Leland Barnard, Ph.D.¹, Andrew Blumenfeld¹, Boris Boutkov¹, Yating Chai, Ph.D.¹, Gisu Eom¹, Lukas Habegger, Ph.D.¹, Young Hahn¹, Alicia Hawes¹, Shareef Khalid¹, Olga Krasheninina¹, Rouel Lanche¹, Adam Mansfield¹, Evan K. Maxwell, Ph.D.¹, Mona Nafde¹, Sean O’Keeffe, Ph.D.¹, John Penn¹, Ayesha Rasool¹, William Salerno, Ph.D.¹, Jeffrey C. Staples, Ph.D.¹, Jeffrey G. Reid, Ph.D.¹

Contribution: Performed and are responsible for analysis needed to produce exome and genotype data, provided compute infrastructure development and operational support, provided variant and gene annotations and their functional interpretation of variants, and conceived and are responsible for creating, developing, and deploying analysis platforms and computational methods for analyzing genomic data.

Clinical Informatics:

Nilanjana Banerjee, Ph.D.¹, Michael Cantor, M.D.¹, Dadong Li Ph.D.¹, Fabricio Sampaio Peres Kury M.D.¹, Deepika Sharma B.H.M.S.¹, Ashish Yadav¹

Contribution: All authors contributed to the development and validation of clinical phenotypes used to identify study participants and (when applicable) controls.

Analytical Genomics and Data Science:

Goncalo R. Abecasis, D.Phil.¹, Joshua Backman, Ph.D.¹, Mathew Barber, Ph.D.¹, Christian Benner, Ph.D.¹, Shan Chen, Ph.D.¹, Amy Damask, Ph.D.¹, Manuel Allen Revez Ferreira, Ph.D.¹, Lauren Gurski¹, Jack Kosmicki, Ph.D.¹, Alexander Li, Ph.D.¹, Nan Lin, Ph.D.¹, Daren Liu¹, Jonathan Marchini Ph.D.¹, Anthony Marcketta¹, Joelle Mbatchou, Ph.D.¹, Shane McCarthy, Ph.D.¹, Colm O’Dushlaine, Ph.D.¹, Charles Paulding, Ph.D.¹, Claudia Schurmann, Ph.D.¹, Dylan Sun¹, Cristopher Van Hout, Ph.D.¹, Kyoko Watanabe, Ph.D.¹, Bin Ye¹, Andrey Ziyatdinov, Ph.D.¹

Contribution: Development of statistical analysis plans. QC of genotype and phenotype files and generation of analysis ready datasets. Development of

statistical genetics pipelines and tools and use thereof in generation of the association results. QC, review and interpretation of result. Generation and formatting of results for manuscript figures.

Therapeutic Area Genetics:

Ariane Ayer¹, Giovanni Coppola M.D.¹, Silvio Alessandro Di Gioia, Ph.D.¹, Jan Freudenberg, M.D.¹, Sahar Gelfman, Ph.D.¹, Claudia Gonzaga-Jauregui, Ph.D.¹, Nehal Gosalia, Ph.D.¹, Julie Horowitz, Ph.D.¹, Luca Lotta M.D. Ph.D.¹, Kavita Praveen, Ph.D.¹

Contribution: Development of study design and analysis plans. Development and QC of phenotype definitions. QC, review, and interpretation of association results.

Functional Modeling:

Shek Man Chim, Ph.D.¹, Giusy Della Gatta, Ph.D.¹, Aris Economides, Ph.D.¹, Lawrence Miloscio¹, Harikiran Nistala, Ph.D.¹, Trikaldarshi Persaud¹

Contribution: Development of *in vivo* and *in vitro* experimental biology and interpretation.

Planning, Strategy, and Operations:

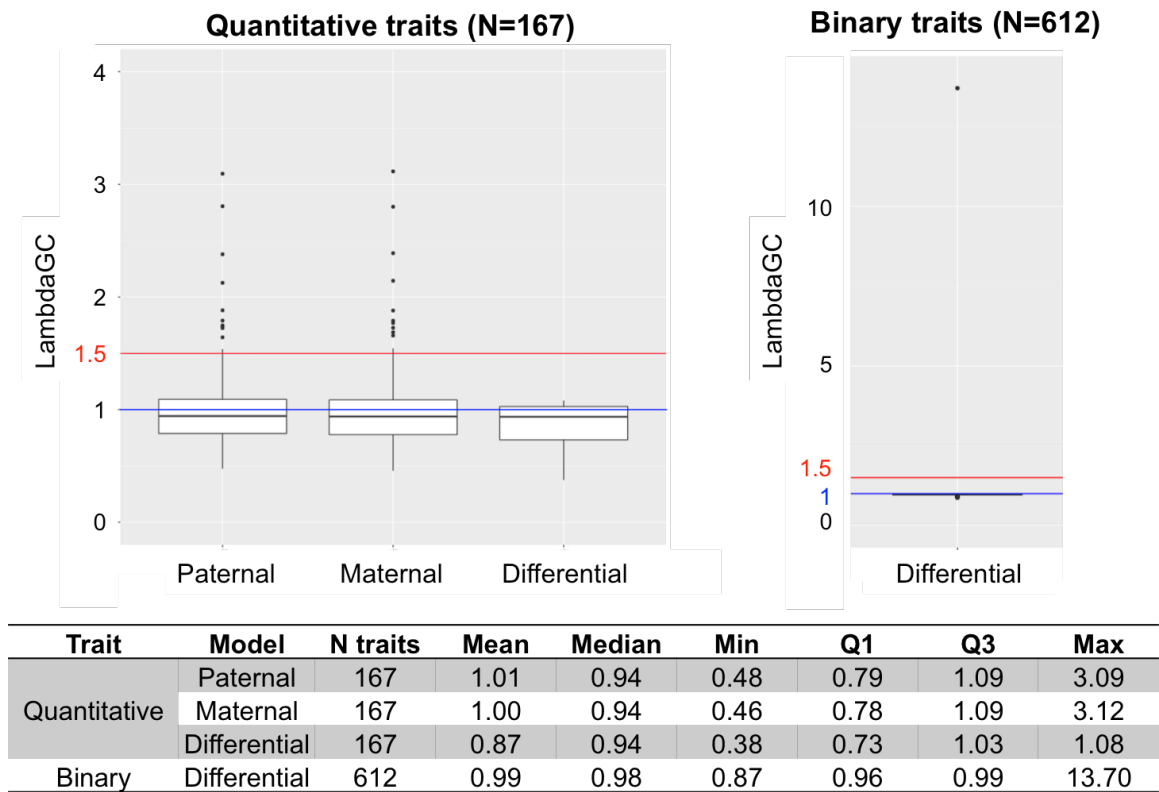
Paloma M. Guzzardo, Ph.D.², Marcus B. Jones, Ph.D.², Michelle LeBlanc, Ph.D.², Jason Mighty, Ph.D.², Lyndon J. Mitnaul, Ph.D.²

Contribution: Contributed to the management and coordination of all research activities, planning and execution, managed the review of the project.

Affiliations:

¹ Regeneron Genetics Center, Tarrytown, NY USA, ² Regeneron Pharmaceuticals, Tarrytown, NY USA

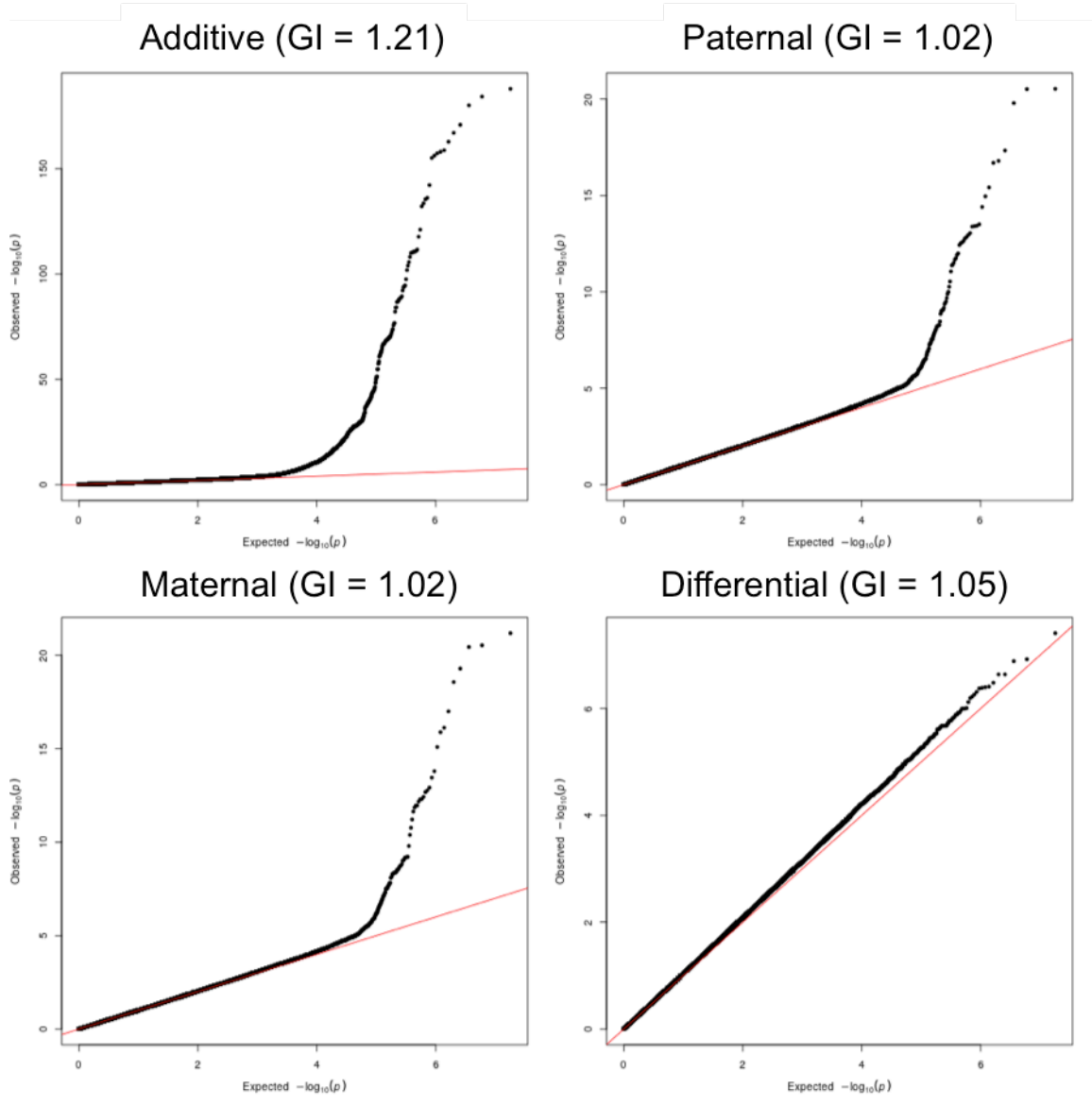
Figure S1. Distribution of genomic inflation factors under PoO models



Distribution of lambdaGC values from GWAS results across 167 quantitative and 612 binary traits under different PoO statistical models is shown. 13 quantitative and 1 binary traits with genomic inflation >1.5 under any statistical model were omitted from further analyses.

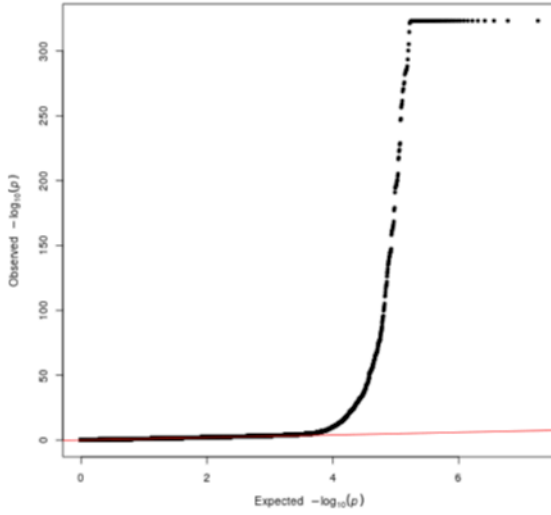
Figure S2. Quantile-Quantile plots under each statistical model for 6 traits with significant and replicated PoO associations

% Monocyte

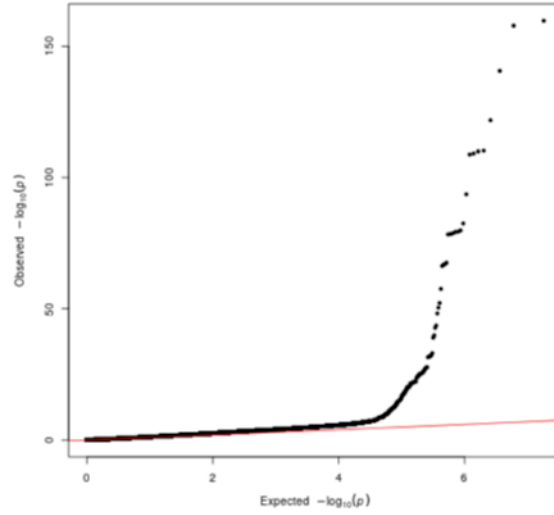


Bilirubin

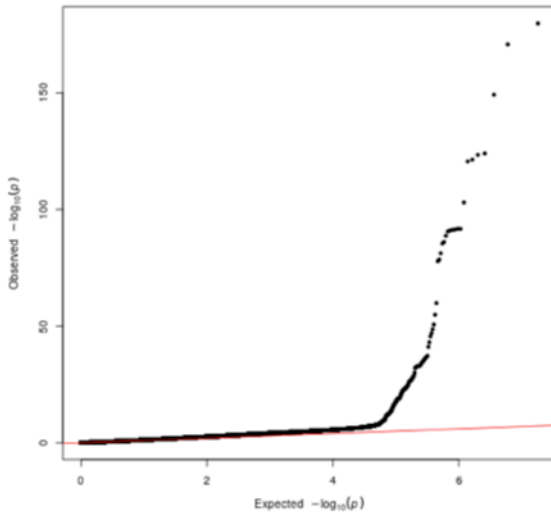
Additive (GI = 1.26)



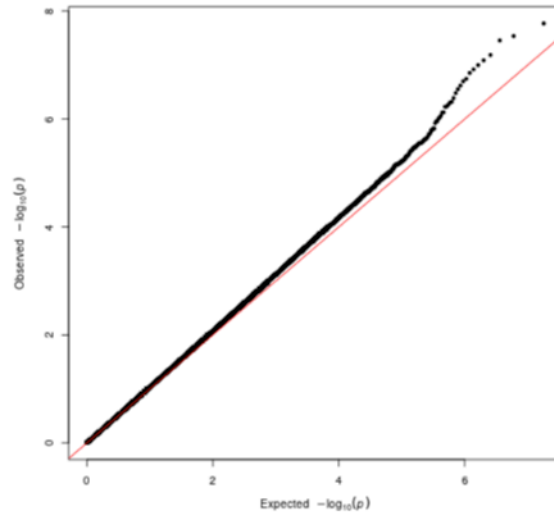
Paternal (GI = 1.36)



Maternal (GI = 1.37)

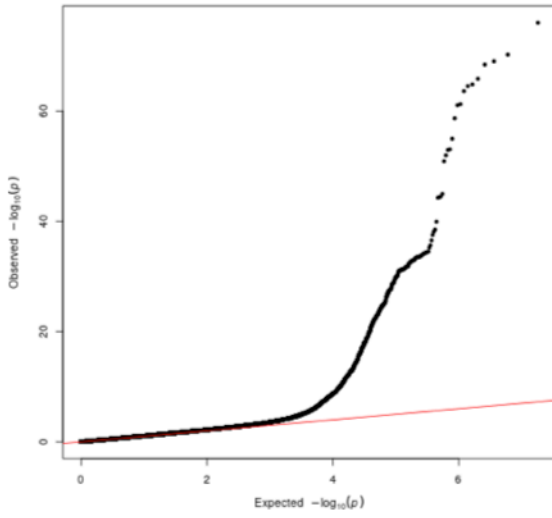


Differential (GI = 1.04)

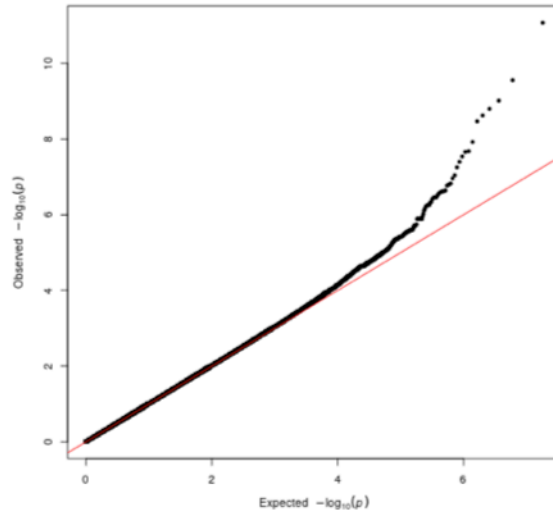


Protein

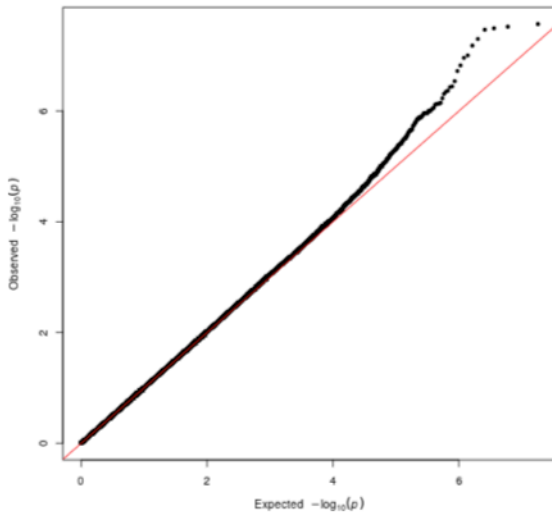
Additive (GI = 1.16)



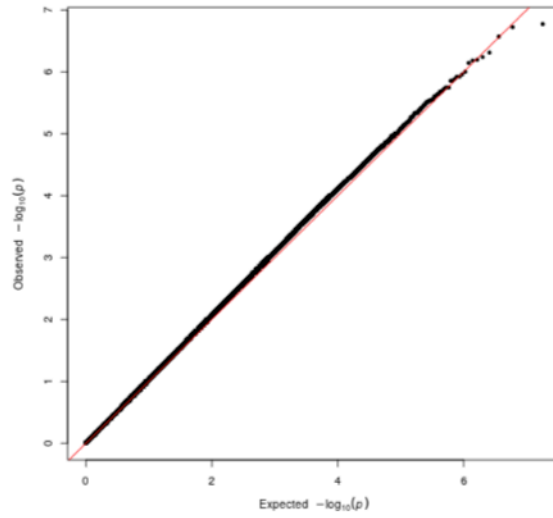
Paternal (GI = 1.01)



Maternal (GI = 1.01)

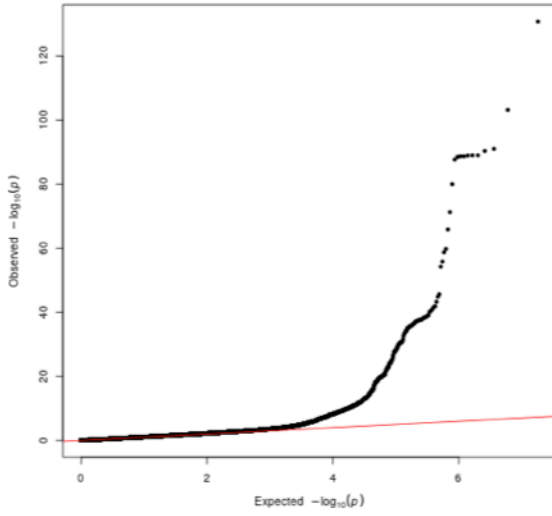


Differential (GI = 1.04)

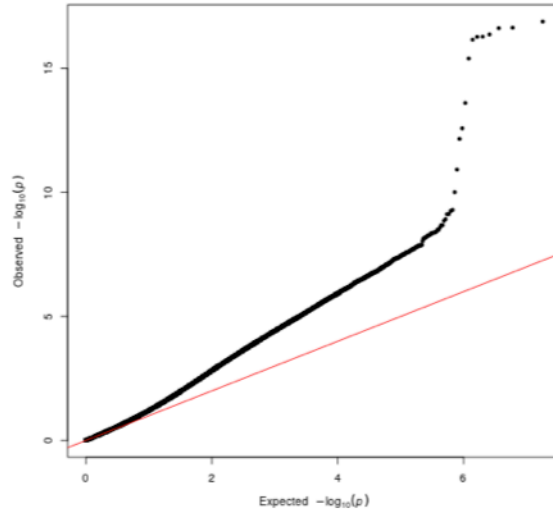


Red blood cell counts

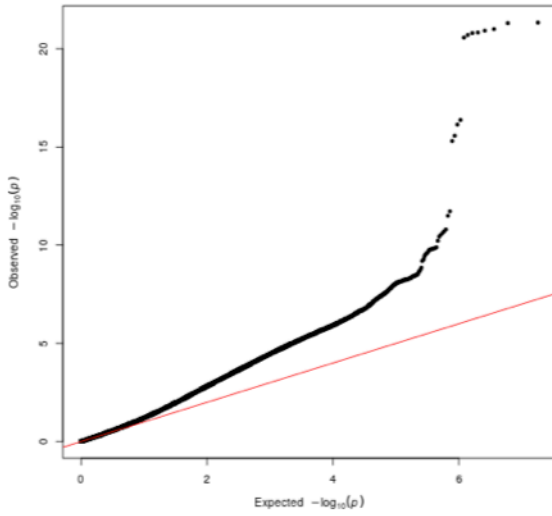
Additive (GI = 1.16)



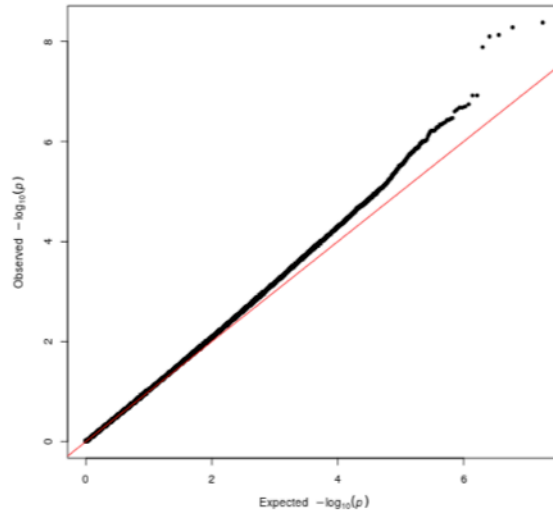
Paternal (GI = 1.39)



Maternal (GI = 1.40)

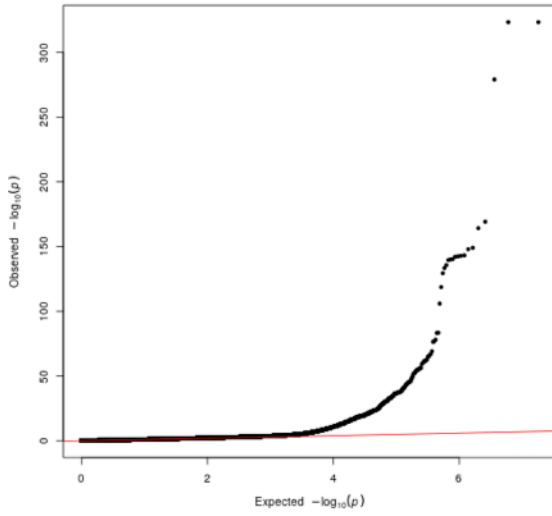


Differential (GI = 1.05)

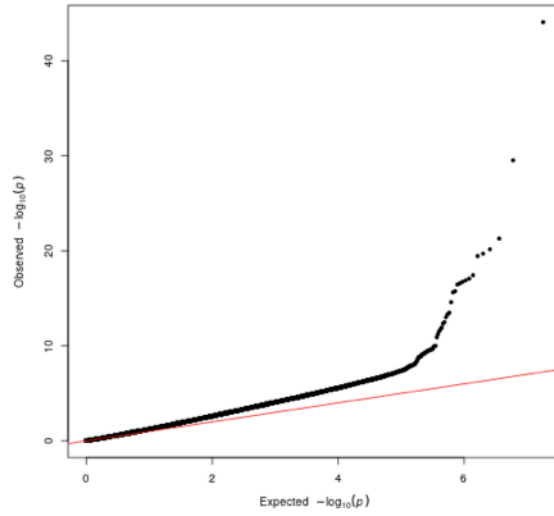


HDL cholesterol

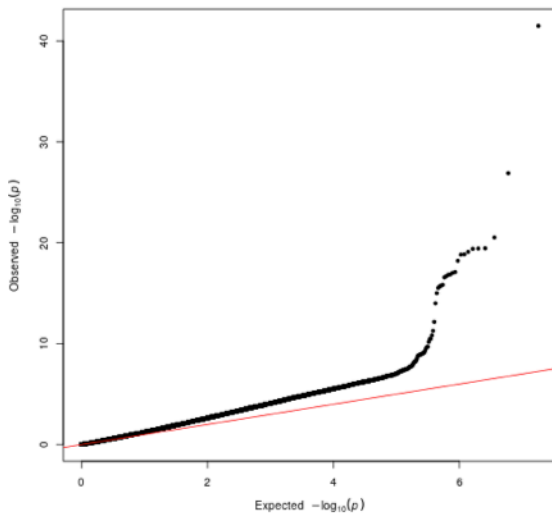
Additive (GI = 1.22)



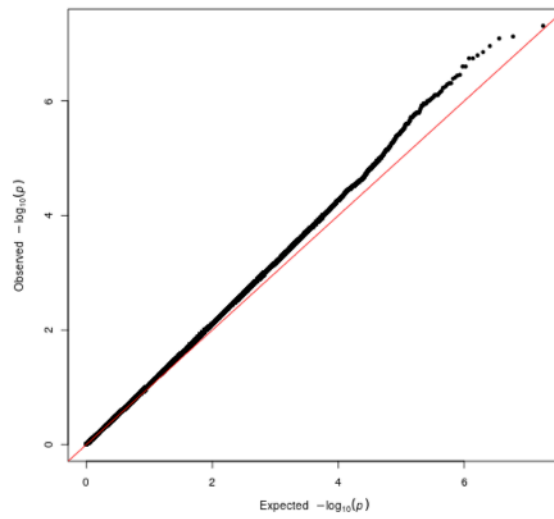
Paternal (GI = 1.35)



Maternal (GI = 1.36)

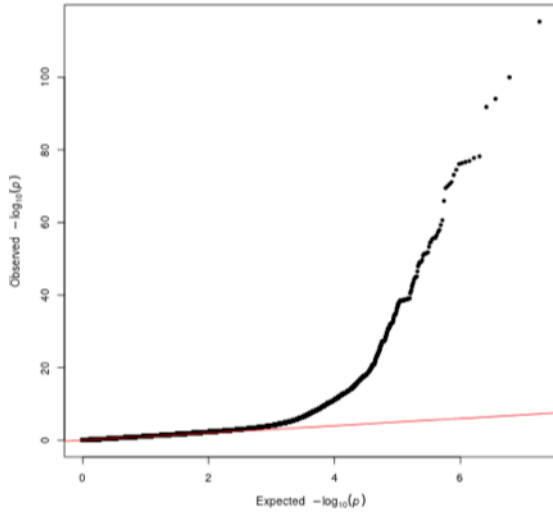


Differential (GI = 1.07)

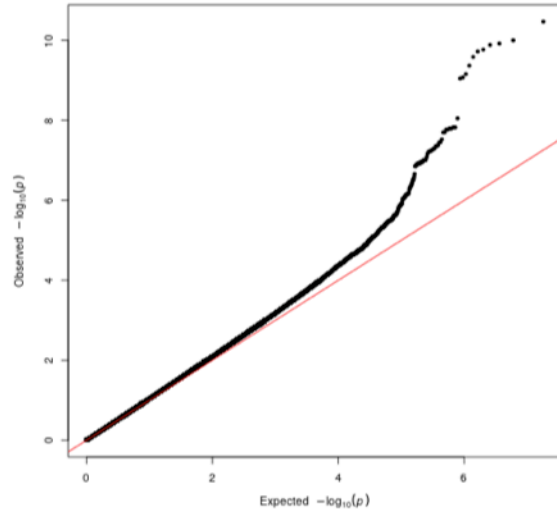


Platelets

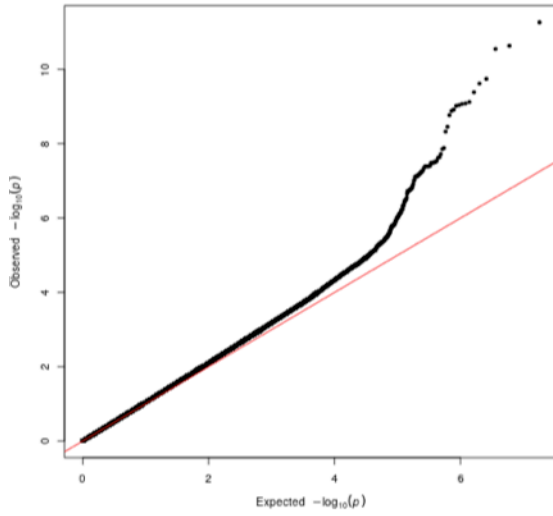
Additive (GI = 1.24)



Paternal (GI = 1.06)



Maternal (GI = 1.05)



Differential (GI = 1.06)

