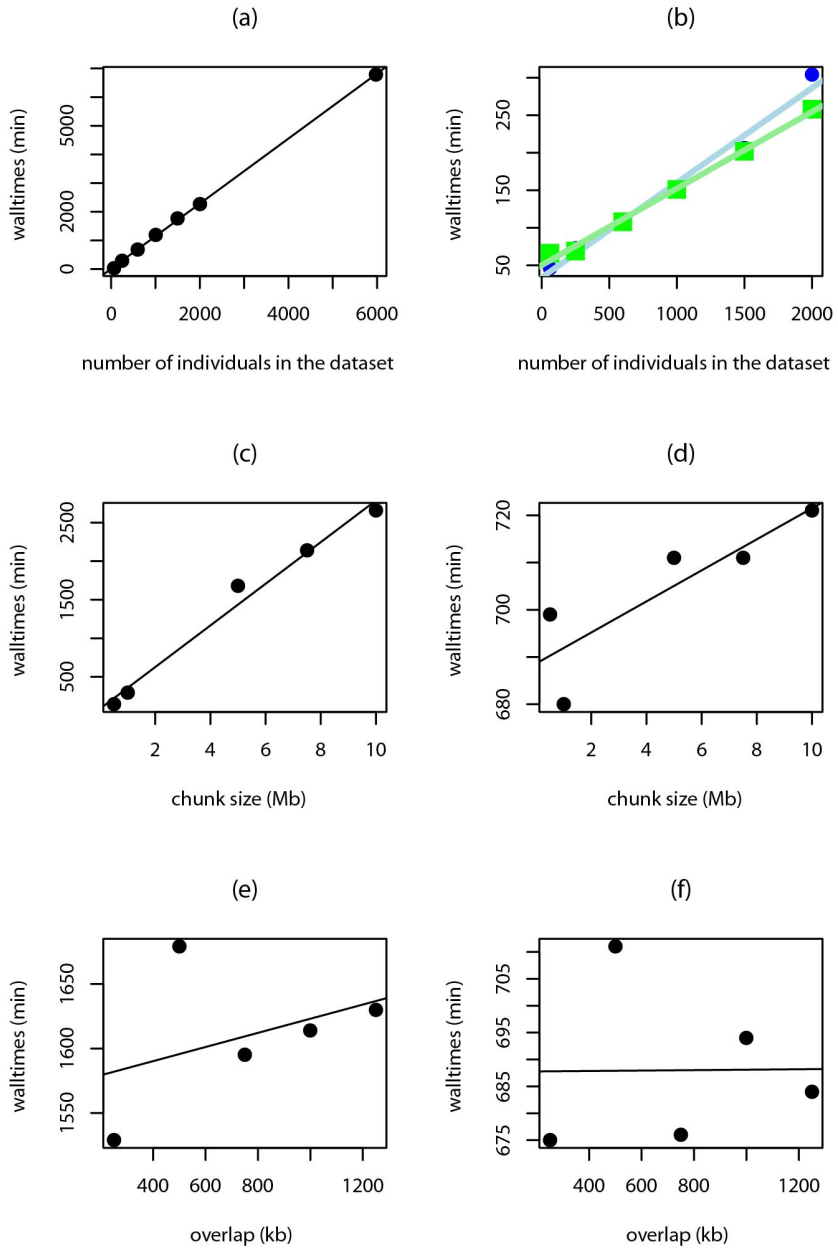

Supplementary Figure 1

The walltimes when splitting up the data set.

The walltimes per job for MaCH (a, c, e) and minimac (b, d, f) for various ways of splitting up the data set. The walltime is the time as measured by a clock on the wall (CPU time, disk writing etcetera) required to impute the target set. The walltime per job for running MaCH fits the linear regression models *t=8.6 + 1.13n* (Figure a), *t=86.49 + 270.02n* (Figure c) and *t=1568.3 + 2.7n* (Figure e). The walltime per job for running minimac fits the linear regression model *t=33.8 + 0.13n* (split before MaCH (blue circles)), *t=50.2 + 0.10n* (split after MaCH (green squares)) (Figure b), *t=688.6 + 3.29n* (Figure d) and *t=687.7 + 0.02n* (Figure f). *t* is the walltime in minutes and *n* the number of samples (a, b), the size of the chunks in Mb (c, d) and the percentage of overlap (e, f). The percentage overlap is 10% in Figure c and d and the chunk size is 5Mb in Figure e and f.

Supplementary Note: the Genome of the Netherlands consortium members:

Analysis group: Morris A. Swertz[6,7] (Co-Chair), Laurent C. Francioli[1], Freerk van Dijk[6,7], Androniki Menelaou[1], Pieter B.T. Neerincx[6,7], Sara L. Pulit[1], Patrick Deelen[6,7], Clara C. Elbers[1], Pier Francesco Palamara[2], Itsik Pe'er[2,8], Abdel Abdellaoui[9], Wigard P. Kloosterman[1], Mannis van Oven[10], Martijn Vermaat[11], Mingkun Li[12], Jeroen F.J. Laros[11], Mark Stoneking[12], Peter de Knijff[13], Manfred Kayser[10], Jan H. Veldink[14], Leonard H. van den Berg[14], Heorhiy Byelas[6,7], Johan T. den Dunnen[11], Martijn Dijkstra[6,7], Najaf Amin[15], K. Joeri van der Velde[6,7], Jouke Jan Hottenga[9], Jessica van Setten[1], Elisabeth M. van Leeuwen[15], Alexandros Kanterakis[6,7], Mathijs Kattenberg[9], Lennart C. Karssen[15], Barbera D.C. van Schaik[16], Jan Bot17, Isaäc J. Nijman[1], David van Enckevort[18], Hailiang Mei[18], Vyacheslav Koval[19], Kai Ye[20,21], Eric-Wubbo Lameijer[21], Matthijs H. Moed[21], Jayne Y. Hehir-Kwa[22], Robert E. Handsaker[5,23], Shamil R. Sunyaev[4,5], Mashaal Sohail[4,5], Fereydoun Hormozdiari[24], Tobias Marschall[25], Alexander Schönhuth[25], Victor Guryev[26], Paul I.W. de Bakker[1,3-5] (Co-Chair);

Cohort collection and sample management group: P. Eline Slagboom[21], Marian B. Beekman[21], Anton J.M. de Craen[21], H. Eka D. Suchiman[21], Albert Hofman[15], Cornelia van Duijn[15], Dorret I. Boomsma[9], Gonneke Willemsen[9], Bruce H. Wolffenbuttel[27], Mathieu Platteel[6], Steven J. Pitts[28], Shobha Potluri[28], David R. Cox[28,34];

Whole-genome sequencing: Qibin Li[29], Yingrui Li[29], Yuanping Du[29], Ruoyan Chen[29], Hongzhi Cao[29], Ning Li[30], Sujie Cao[30], Jun Wang[29,31,32];

Ethical, Legal, and Social Issues: Jasper A. Bovenberg[33];

Steering committee: Cisca Wijmenga[6,7] (Principal Investigator), Morris A. Swertz[6,7], Cornelia M. van Duijn[15], Dorret I. Boomsma[9], P. Eline Slagboom[21], Gertjan B. van Ommen[11], Paul I.W. de Bakker[1,3-5]

[1] Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands. [2] Department of Computer Science, Columbia University, New York, NY, USA. [3] Department of Epidemiology, University Medical Center Utrecht, Utrecht, The Netherlands. [4] Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [5] Broad Institute of Harvard and MIT, Cambridge, MA, USA. [6] Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [7] Genomics Coordination Center, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [8] Department of Systems Biology, Columbia University, New York, NY, USA. [9] Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands. [10] Department of Forensic Molecular Biology, Erasmus Medical Center, Rotterdam, The Netherlands. [11] Leiden Genome Technology Center, Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands. [12] Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. [13] Forensic Laboratory for DNA Research, Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands. [14] Department of Neurology, University Medical Center Utrecht, Utrecht, The Netherlands. [15] Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. [16] Bioinformatics Laboratory, Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam Medical Center, Amsterdam, The Netherlands. [17] SURFsara, Science Park, Amsterdam, The Netherlands. [18] Netherlands Bioinformatics Centre, Nijmegen, The Netherlands. [19] Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. [20] The Genome Institute, Washington University, St. Louis, MO, USA. [21] Section of Molecular Epidemiology, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands. [22] Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [23] Department of Genetics, Harvard Medical School, Boston, MA, USA. [24] Department of Genome Sciences, University of Washington, Seattle, WA, USA. [25] Centrum voor Wiskunde en Informatica, Life Sciences Group, Amsterdam, The Netherlands. [26] European Research Institute for the Biology of Ageing, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [27] Department of Endocrinology, University Medical Center Groningen, Groningen, The Netherlands. [28] Rinat-Pfizer Inc, South San Francisco, CA, USA. [29] BGI-Shenzhen, Shenzhen, China. [30] BGI-Europe, Copenhagen, Denmark. [31] Department of Biology, University of Copenhagen, Copenhagen, Denmark. [32] The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark. [33] Legal Pathways Institute for Health and Bio Law, Aerdenhout, The Netherlands. [34] Deceased.