# Linkage disequilibrium and haplotype homozygosity in population samples genotyped at a high marker density

Hui Wang[1], Chia-Ho Lin[1], Susan Service[2], The international collaborative group on isolated populations, Yuguo Chen[3], Nelson Freimer[2], and Chiara Sabatti[1,4]

Department of Statistics[1], Center for Neurobehavioral Genetics[2], and Department of Human Genetics[4], UCLA, Los Angeles CA 90095-7088. Department of Statistics[3], University of Illinois at Urbana-Champain, Champain IL 61802.

**Supplementary material**

**Corresponding author**  Chiara Sabatti

Department of Human Genetics

UCLA School of Medicine

695 Charles E. Young Drive South

Los Angeles, California 90095-7088 (USA)

FAX: (310) 794-5446

Phone: (310) 794-9567

e-mail: csabatti@mednet.ucla.edu

**The international collaborative group on isolated populations**  members that are not listed separately as authors of this manuscript are: Maria Karayiorgou[a], J. Louw Roos[b], Herman Pretorious[b], Gabriel Bedoya[c], Jorge Ospina[d], Andres Ruiz-Linares[c,e], António Macedo[f], Joana Almeida Palha[g], Peter Heutink[h,i], Yurii Aulchenko[j], Ben Oostra[j], Cornelia van Duijn[j], Marjo-Riitta Jarvelin[k,l], Teppo Varilo[m,n], Lynette Peddle[o], Proton Rahman[p], Giovanna Piras[q], Maria Monne[q], Leena Peltonen[m,n].

[a]Rockefeller University, New York, New York, USA; [b]University of Pretoria Weskoppies Hospital, Pretoria, Republic of South Africa; [c]Laboratorio de Genetica Molecular, Universidad de Antioquia, Medellin, Colombia; [d]Departamento de Psiquiatria, Universidad de Antioquia, Medellin, Colombia; [e]The Galton Laboratory, Department of Biology (Wolfson House), University College London, London, UK; [f]Instituto de Psicologia Médica, Faculdade de Medicina, Coimbra, Portugal; [g]Life and Health Sciences Research Institute

(ICVS), School of Health Sciences, University of Minho, Braga, Portugal; [h]Department of Human Genetics, Section of Medical Genomics, VU University Medical Center and [i]Center for Neurogenomics and Cognitive Research, VU University and VU University Medical Center, 1081 BT, Amsterdam, The Netherlands; [j]Genetic Epidemiology Unit, Departments of Epidemiology  Biostatistics and Clinical Genetics, Erasmus MC, PO BOX 1738, 3000 DR Rotterdam, The Netherlands; [k]Department of Public Health Science and General Practice, University of Oulu, Oulu, Finland; [l]Department of Epidemiology and Public Health, Imperial College London, London, UK; [m]Department of Medical Genetics, University of Helsinki, Helsinki, Finland; [n]Department of Molecular Medicine, National Public Health Insitute, Biomedicum Helsinki, 00290 Helsinki, Finland; [o]Newfound Genomics, Newfoundland, Canada; [p]Memorial University of Newfoundland, Newfoundland, Canada; [q]Division of Haematology, San Francesco Hospital, ASL3, Nuoro Italy;

|     | ant | ash | azo | cau | cr | erf | fic | fik | fip | nfl | saf | sar |
|-----|-----|-----|-----|-----|----|-----|-----|-----|-----|-----|-----|-----|
| ant | *   | 0.63 | 11.878 | 5.256 | 11.421 | 0.299 | >15.66 | >15.66 | 5.394 | 1.934 | 4.319 | 2.980 |
| ash | --  | *   | 15.955 | 9.388 | 7.189 | 0.339 | >15.66 | >15.66 | 8.310 | 0.390 | 2.137 | 2.412 |
| azo | --  | --  | *   | >15.66 | >15.66 | 12.932 | >15.66 | >15.66 | 2.347 | >15.66 | >15.66 | >15.66 |
| cau | --  | --  | --  | *   | 0.548 | 13.531 | 5.943 | 7.146 | >15.66 | 5.832 | 2.726 | 7.161 |
| cr  | --  | --  | --  | --  | *  | 11.849 | 5.044 | 3.759 | >15.66 | 4.699 | 1.842 | 5.903 |
| erf | --  | --  | --  | --  | -- | *   | >15.66 | 13.092 | 6.291 | 1.116 | 3.956 | 1.510 |
| fic | --  | --  | --  | --  | -- | --  | *   | 1.566 | >15.66 | 15.955 | 8.565 | 11.709 |
| fik | --  | --  | --  | --  | -- | --  | --  | *   | >15.66 | 12.296 | 6.678 | 7.171 |
| fip | --  | --  | --  | --  | -- | --  | --  | --  | *   | 10.294 | 15.955 | 13.132 |
| nfl | --  | --  | --  | --  | -- | --  | --  | --  | --  | *   | 1.010 | 2.019 |
| saf | --  | --  | --  | --  | -- | --  | --  | --  | --  | --  | *   | 0.951 |
| sar | --  | --  | --  | --  | -- | --  | --  | --  | --  | --  | --  | *   |

Table 1: P-values on negative log 10 scale of a Kolmogorov Smirnov test to compare the distributions of estimated NRH segments lengths across the 12 populations. ">" refers to the tests whose p-values can not be accurately computed due to ties.
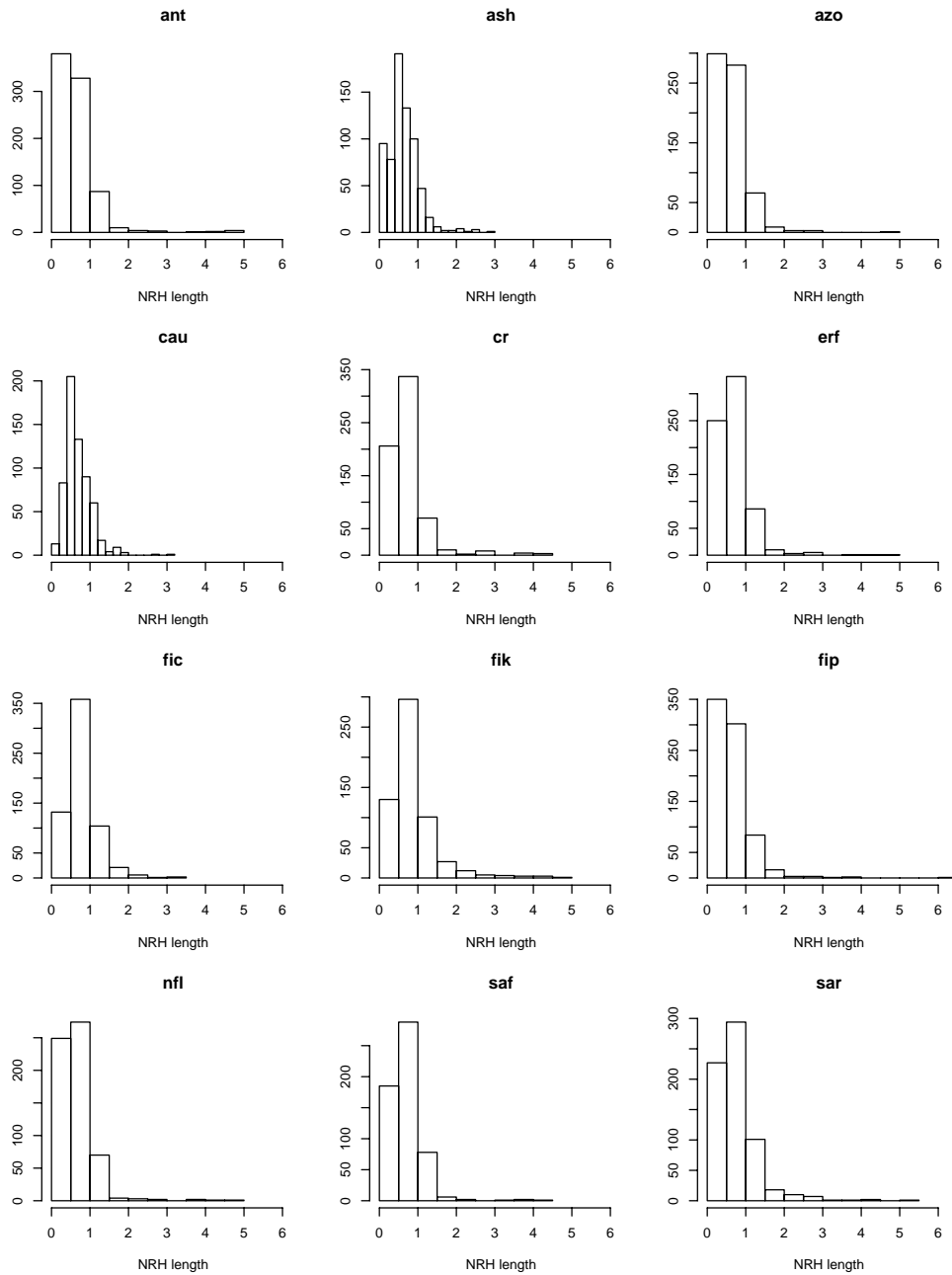
1

Figure 1: **Distribution of estimated NRH segments lengths for the 12 populations**. The segments lengths (in Mb) have been subjected to a square root transformation to reduce skewness.
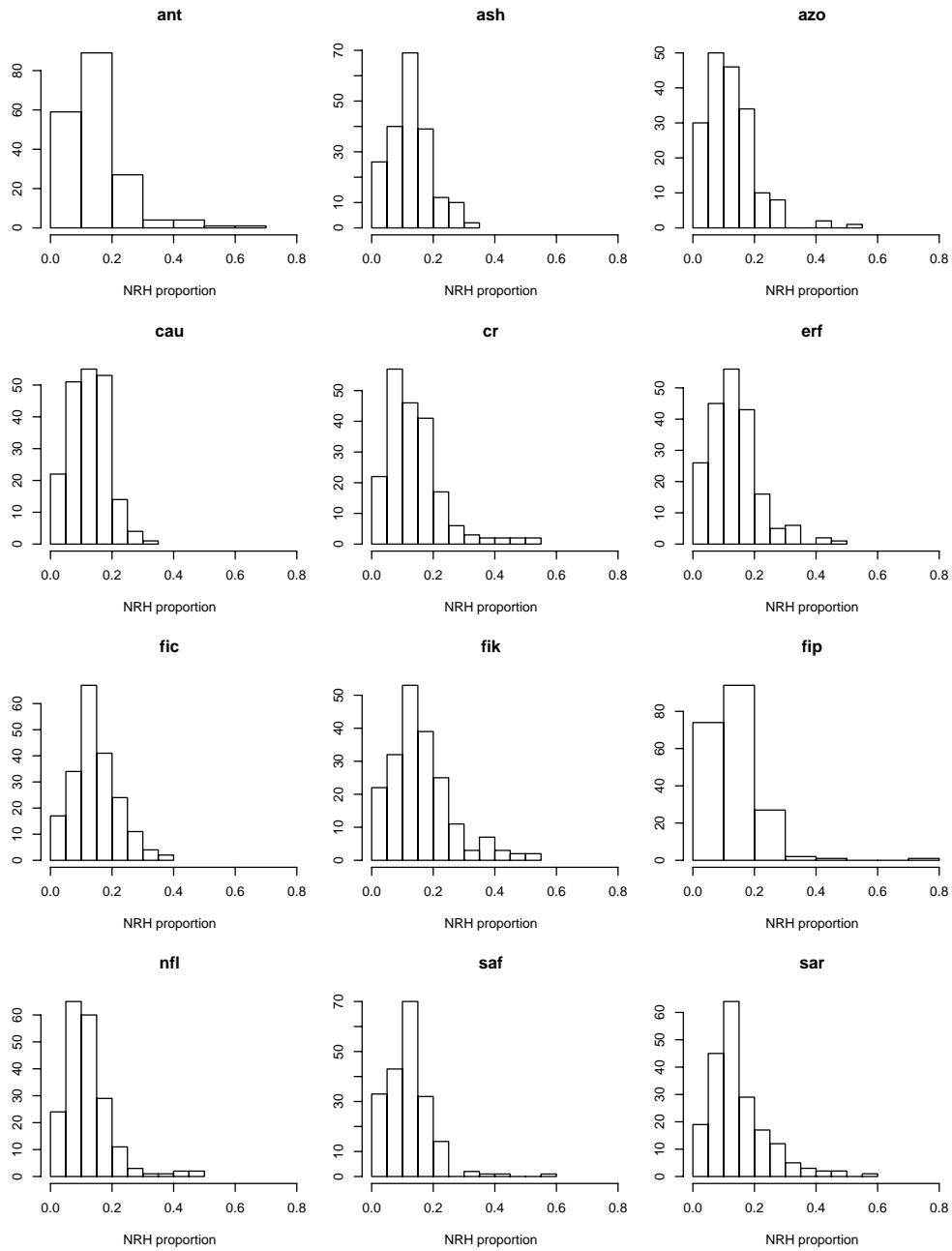
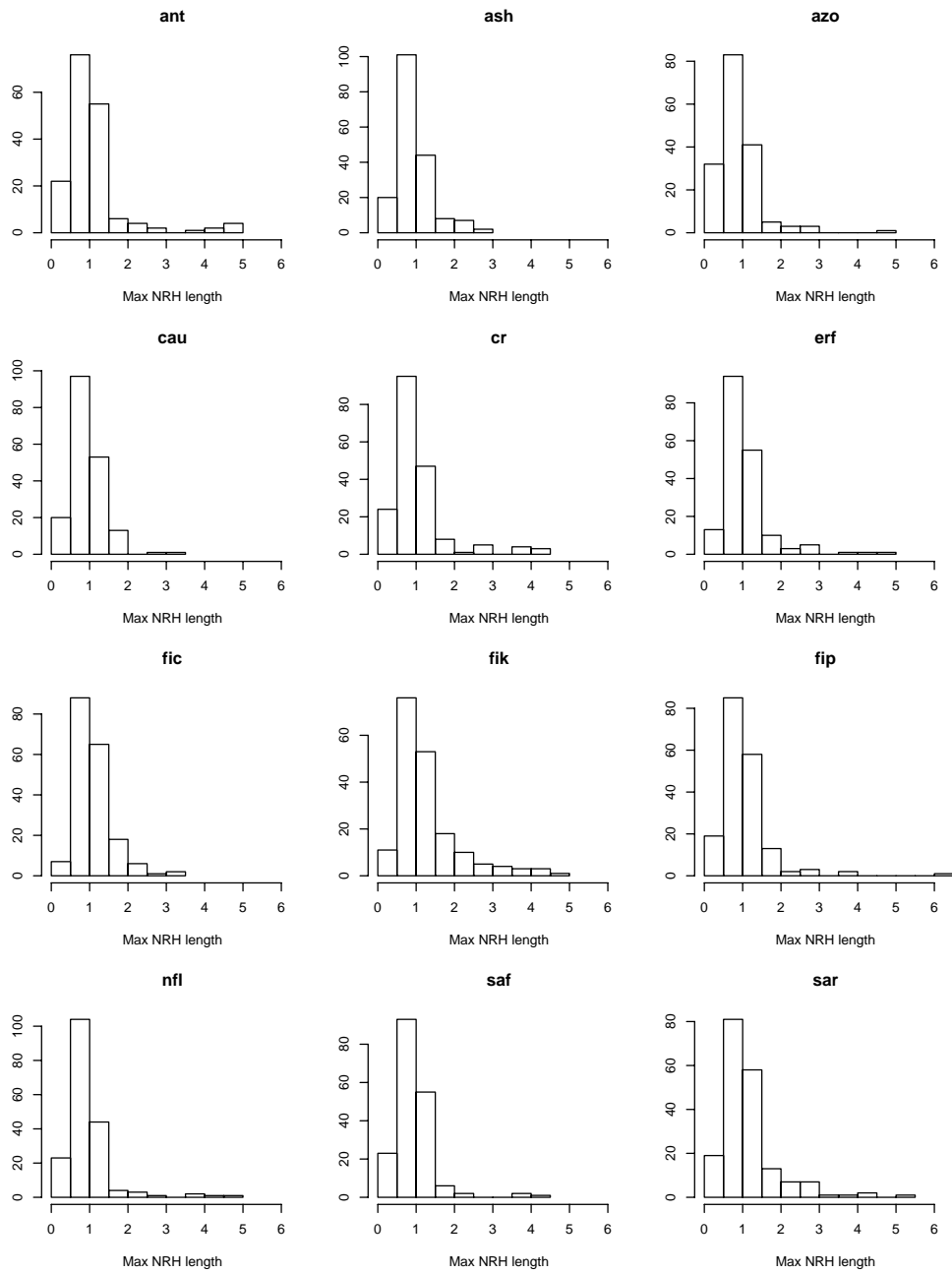Figure 2: **Distribution of chromosome proportion occupied by NRH segments lengths for the 12 populations**.

Figure 3: **Distribution of maximum NRH segments lengths per individual for the 12 popula-tions**. The segments lengths (in Mb) have been subjected to a square root transformation to reduce skewness.