**Supplementary Information**

1. **Formal definition of heteroplasmy change rate**

Heteroplasmy *h(t)* is a function of an initial heteroplasmy *h(t=0)* and time *t*, with describing the rate of change of heteroplasmy:

.

High values correspond to high proliferative differences (in our analysis, *H1* dominating over *H0*), and hence high probabilities of amplification of one mtDNA type with time. This expression has been derived and used in the analysis of segregation in mouse models ([Burgstaller *et al.* , 2014](#_ENREF_3)) and emerges from more detailed mathematical analysis of mtDNA evolution within cells ([Johnston *et al.* , 2015](#_ENREF_9)).

**II. UK city mtDNA haplogroup compositions**

In order to produce a realistic approximation of the mitochondrial haplogroup composition of the two largest UK cities, London and Birmingham, an approach based on ethnic categories in the 2011 UK census and mt haplogroup data from corresponding regions and/or countries was employed.

Ethnic composition data were downloaded from <http://webarchive.nationalarchives.gov.uk>, in which 18 categories were defined (see Table 1 for these category definitions and their frequencies in London and Birmingham).

**Table 1.** Proportion of different census categories in London and Birmingham

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Code | Category | Sub-category | London (%) | B'ham (%) |
| Pop1 | White | English/Welsh/Scottish/Northern Irish/British | 44.8 | 53.1 |
| Pop2 | White | Irish | 2.2 | 2.1 |
| Pop3 | White | Gypsy or Irish Traveller | 0.1 | 0 |
| Pop4 | White | Other White | 12.6 | 2.7 |
| Pop5 | Mixed/multiple ethnic group | White and Black Caribbean | 1.5 | 2.3 |
| Pop6 | Mixed/multiple ethnic group | White and Black African | 0.8 | 0.3 |
| Pop7 | Mixed/multiple ethnic group | White and Asian | 1.2 | 1 |
| Pop8 | Mixed/multiple ethnic group | Other Mixed | 1.4 | 0.8 |
| Pop9 | Asian/Asian British | Indian | 6.6 | 6 |
| Pop10 | Asian/Asian British | Pakistani | 2.7 | 13.5 |
| Pop11 | Asian/Asian British | Bangladeshi | 2.7 | 3 |
| Pop12 | Asian/Asian British | Chinese | 1.5 | 1.2 |
| Pop13 | Asian/Asian British | Other Asian | 4.9 | 2.9 |
| Pop14 | Black/African/Caribbean/Black British | African | 7 | 2.8 |
| Pop15 | Black/African/Caribbean/Black British | Caribbean | 4.2 | 4.4 |
| Pop16 | Black/African/Caribbean/Black British | Other | 2.1 | 1 |
| Pop17 | Other ethnic group | Arab | 1.3 | 1 |
| Pop18 | Other ethnic group | Any other ethnic group | 2.1 | 1 |

Then, for each ethnic category suitable mtDNA datasets were collated (see Table 2 for resultant haplogroup profiles).

**Table 2.** Estimated haplogroup frequencies in UK census categories.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| HG | Pop1 | Pop2 | Pop3 | Pop4 | Pop5 | Pop6 | Pop7 | Pop8 | Pop9 | Pop10 | Pop11 | Pop12 | Pop13 | Pop14 | Pop15 | Pop16 | Pop17 | Pop18 |
| A | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% | 0.0% | 1.4% | n/a | 0.0% | 0.1% | 10.0% | 5.3% | 2.7% | 0.0% | 0.3% | 0.1% | 0.1% | n/a |
| B | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% | 0.0% | 5.2% | n/a | 0.2% | 0.4% | 0.0% | 20.5% | 10.4% | 0.0% | 0.3% | 0.1% | 0.0% | n/a |
| C | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 1.1% | n/a | 0.5% | 0.0% | 0.0% | 3.8% | 2.1% | 0.0% | 0.0% | 0.0% | 0.0% | n/a |
| D | 0.0% | 0.0% | 0.0% | 0.0% | 0.3% | 0.0% | 3.9% | n/a | 0.1% | 0.0% | 0.0% | 15.3% | 7.7% | 0.0% | 0.5% | 0.3% | 0.0% | n/a |
| F | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 4.8% | n/a | 1.1% | 1.2% | 0.0% | 18.0% | 9.6% | 0.0% | 0.0% | 0.0% | 0.0% | n/a |
| G | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.8% | n/a | 0.8% | 0.6% | 0.0% | 2.4% | 1.6% | 0.0% | 0.0% | 0.0% | 0.0% | n/a |
| H | 45.2% | 50.7% | 50.7% | 48.1% | 22.7% | 24.4% | 23.2% | n/a | 2.5% | 11.2% | 0.0% | 0.0% | 1.3% | 3.6% | 0.3% | 1.9% | 24.1% | n/a |
| I | 4.1% | 3.0% | 3.0% | 2.9% | 2.0% | 2.4% | 2.1% | n/a | 0.2% | 0.7% | 0.0% | 0.0% | 0.1% | 0.8% | 0.0% | 0.4% | 1.4% | n/a |
| J | 12.4% | 11.0% | 11.0% | 9.2% | 6.4% | 6.5% | 6.5% | n/a | 1.5% | 3.5% | 0.0% | 0.0% | 0.7% | 0.6% | 0.5% | 0.6% | 11.2% | n/a |
| K | 8.3% | 12.3% | 12.3% | 5.2% | 4.1% | 4.4% | 4.3% | n/a | 0.7% | 1.1% | 0.0% | 0.0% | 0.3% | 0.5% | 0.0% | 0.2% | 5.4% | n/a |
| L0 | 0.0% | 0.0% | 0.0% | 0.0% | 2.5% | 5.8% | 0.0% | n/a | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% | 11.6% | 5.0% | 8.3% | 2.1% | n/a |
| L1 | 0.0% | 0.0% | 0.0% | 0.0% | 9.8% | 7.2% | 0.0% | n/a | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 14.3% | 19.5% | 17.0% | 1.9% | n/a |
| L2 | 0.0% | 0.0% | 0.0% | 0.0% | 19.4% | 15.1% | 0.0% | n/a | 0.0% | 0.5% | 0.0% | 0.0% | 0.0% | 29.9% | 38.8% | 34.5% | 4.7% | n/a |
| L3 | 0.1% | 0.0% | 0.0% | 0.0% | 16.5% | 14.1% | 0.0% | n/a | 0.0% | 1.3% | 0.0% | 0.0% | 0.0% | 27.8% | 33.0% | 30.5% | 7.8% | n/a |
| M | 0.0% | 0.0% | 0.0% | 0.0% | 0.4% | 1.8% | 20.5% | n/a | 54.3% | 38.1% | 66.7% | 27.6% | 40.9% | 3.5% | 0.8% | 2.1% | 4.0% | n/a |
| N | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% | 0.3% | n/a | 0.1% | 1.3% | 0.0% | 1.2% | 0.6% | 0.1% | 0.0% | 0.1% | 4.3% | n/a |
| O | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | n/a | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | n/a |
| P | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | n/a | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | n/a |
| R | 0.1% | 0.3% | 0.3% | 0.0% | 0.1% | 0.2% | 5.7% | n/a | 18.0% | 12.5% | 6.7% | 4.4% | 11.2% | 0.3% | 0.0% | 0.1% | 10.4% | n/a |
| S | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | n/a | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | n/a |
| T | 10.5% | 7.0% | 7.0% | 9.7% | 5.2% | 5.7% | 5.6% | n/a | 1.3% | 3.1% | 0.0% | 0.0% | 0.6% | 0.8% | 0.0% | 0.4% | 7.2% | n/a |
| U | 12.6% | 12.3% | 12.3% | 17.0% | 6.8% | 7.5% | 10.1% | n/a | 15.4% | 20.5% | 16.7% | 0.0% | 7.7% | 2.3% | 1.0% | 1.7% | 10.3% | n/a |
| V | 3.2% | 0.0% | 0.0% | 0.1% | 1.6% | 1.7% | 1.6% | n/a | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% | 0.0% | 0.1% | 1.1% | n/a |
| W | 1.5% | 0.0% | 0.0% | 0.2% | 0.7% | 0.8% | 1.5% | n/a | 3.2% | 3.3% | 0.0% | 0.0% | 1.6% | 0.2% | 0.0% | 0.1% | 1.0% | n/a |
| X | 1.8% | 0.7% | 0.7% | 2.0% | 0.9% | 1.0% | 0.9% | n/a | 0.0% | 0.2% | 0.0% | 0.0% | 0.0% | 0.2% | 0.0% | 0.1% | 1.7% | n/a |
| other | 0.3% | 2.7% | 2.7% | 5.7% | 0.3% | 1.4% | 0.6% | n/a | 0.2% | 0.0% | 0.0% | 1.5% | 0.8% | 2.5% | 0.3% | 1.4% | 1.2% | n/a |

**Pop1 White: English/Welsh/Scottish/Northern Irish/British**

These are frequencies from majority ethnic Britons, largely from rural areas, derived from over 4600 individuals ([Helgason *et al.* , 2001](#_ENREF_8), [Røyrvik, unpublished](#_ENREF_17)).

**Pop2 White: Irish**

Data from 300 individuals, from ([Achilli *et al.* , 2007](#_ENREF_1)), originially from ([McEvoy *et al.* , 2004](#_ENREF_11)).

**Pop3 White: Gypsy or Irish Traveller**

Same as Pop3, given the lack of qualitative genetic differentiation between Traveller and non-Traveller Irish ([Relethford and Crawford, 2013](#_ENREF_16)). Note that this category does not include Roma.

**Pop4 White: Other White**

These frequencies are composed of 50% pan-European average frequencies and 50% Polish frequencies, to take account of the large Polish communities in London (see <http://www.migrationobservatory.ox.ac.uk/briefings/migrants-uk-overview> )

and Birmingham, where Poles are 29% of this category. Variations in upper-hierarchy haplogroup frequencies are comparatively minor within Europe, and the New World white populations – derived as they primarily are from European parental populations – deviate little from European averages. Data from ([Achilli *et al.*, 2007](#_ENREF_1)), comprising 10 970 individuals.

**(For ‘Mixed/multiple ethnic group’ categories, see below.)**

**Pop9 Asian/Asian British: Indian**

This is a weighted composite of different frequencies from 2126 individuals from the Indian subcontinent

([Dubut *et al.* , 2009](#_ENREF_6)), based on major regions from which immigrants to the UK originated

([Chanda and Ghosh, 2013](#_ENREF_4)).

Weighting was as follows: Punjab – 45%, Kerala, Tamil Nadu, Andra Pradesh and Gujarat – 13.75% each.

**Pop10 Asian/Asian British: Pakistani**

This unweighted composite of 818 individuals includes all Pakistani groups from ([Quintana-Murci *et al.* , 2004](#_ENREF_15)), as well as Indian Punjabis and Kashmiris from

([Dubut *et al.*, 2009](#_ENREF_6)), as people hereditarily from the latter two geographic regions constitute by far the largest Pakistani sub-populations in the UK

([TheChangeInstitute, 2009](#_ENREF_20)).

**Pop11 Asian/Asian British: Bangladeshi**

From 30 individuals from ([Dubut *et al.*, 2009](#_ENREF_6)).

**Pop12 Asian/Asian British: Chinese**

These frequencies are from 2398 southern Chinese individuals from ([Xue *et al.* , 2008](#_ENREF_21)), as the majority of immigration has traditionally been from southern China ([Pharaoh, 2009](#_ENREF_13)). While this has perhaps changed ([Pharaoh, 2009](#_ENREF_13)), for the purposes of our study, the differences in haplogroup composition, at our level of refinement, would not change materially.

**Pop13 Asian/Asian British: Other Asian**

Defined as 50% Pop9, 50% Pop12. While the composition/ethnicity of different ‘other’ Asian groups is unknown, the Chinese and Indian samples, taken together, should for our purposes adequately represent the majority of East Asia, South Asia and Southeast Asia. Central Asia, with its higher frequencies of typically `West Eurasian' haplogroups, may be underrepresented, but at a small fraction of 4.9-2.9%, this will not substantially affect our calculations, especially as these haplogroups are well represented in Pops1-4.

**Pop14 Black/African/Caribbean/Black British: African**

This is a weighted composite of different frequencies from the African continent, scaled as 55% Western and Central African, 41% South and East African, and 3% North African (see ([Owen, unpublished](#_ENREF_12)) for details on the origins of African immigrants to the UK). Sources for African regional and national haplogroup frequencies are

([Fendt *et al.* , 2012](#_ENREF_7), [Kivisild *et al.* , 2004](#_ENREF_10), [Plaza *et al.* , 2003](#_ENREF_14), [Salas *et al.* , 2002](#_ENREF_18), [Saunier *et al.* , 2009](#_ENREF_19)).

Note that North Africans are also included in pop17.

**Pop15 Black/African/Caribbean/Black British: Caribbean**

From ([Deason *et al.* , 2012](#_ENREF_5)), 400 individuals representing Jamaica. The vast majority of haplogroups are typical of Sub-Saharan Africa, but a small minority are likely the result of European (H, J, U) and Taino (A, B, D) ancestry.

**Pop16 Black/African/Caribbean/Black British: Other**

These frequencies are composed of 50% pop14 and 50% pop15.

**Pop 17 Other ethnic group: Arab**

Base frequencies from ([Badro *et al.* , 2013](#_ENREF_2)) for 3248 individuals. Haplogroup frequencies are weighted averages of 33.3% each from the Arabian Peninsula, the Middle East, and Africa north of the Sahara.

**Pop18 Other ethnic group: Other**

As there is no basis for estimating haplogroup frequencies for this category (2.1% in London, 1.0% in Birmingham), it has not been included in the analysis.

**Pop5 Mixed/multiple ethnic group: White and Black Caribbean**

These frequencies are composed of 50% Pop1 and 50% Pop15. This assumes that the census respondents’ mothers are equally likely to be ‘White’ or ‘Black Caribbean’, which may not be the case.

**Pop6 Mixed/multiple ethnic group: White and Black African**

These frequencies are composed of 50% Pop1 and 50% Pop14. This assumes that the census respondents’ mothers are equally likely to be ‘White’ or ‘Black African’, which may not be the case.

**Pop7 Mixed/multiple ethnic group: White and Asian**

These frequencies are composed of 50% Pop1 and 50% Pop13. This assumes that the proportion of census respondents’ mothers are equally likely to be ‘White’ or ‘Asian’, which may not be the case.

**Pop8 Mixed/multiple ethnic group: Other Mixed**

As there is no basis for estimating haplogroup frequencies for this category (1.4% in London, 0.8% in Birmingham), it has not been included in the analysis.

Note that minority haplogroups from any given sample population are often not assayed in the original source, so the ‘other’ category by necessity has some small overlap with the other haplogroups in the table for different groups. The maximum proportion of ‘other' haplogroups is just 1.27% (in London), so it is not expected to bias the overall dataset unduly.

**References:**

Achilli A, Olivieri A, Pala M, Metspalu E, Fornarino S, Battaglia V, Accetturo M, Kutuev I, Khusnutdinova E, Pennarun E *et al.* (2007) Mitochondrial DNA variation of modern Tuscans supports the near eastern origin of Etruscans. American journal of human genetics 80, 759-68.

Badro DA, Douaihy B, Haber M, Youhanna SC, Salloum A, Ghassibe-Sabbagh M, Johnsrud B, Khazen G, Matisoo-Smith E, Soria-Hernanz DF *et al.* (2013) Y-chromosome and mtDNA genetics reveal significant contrasts in affinities of modern Middle Eastern populations with European and African populations. PloS one 8, e54616.

Burgstaller JP, Johnston IG, Jones NS, Albrechtova J, Kolbe T, Vogl C, Futschik A, Mayrhofer C, Klein D, Sabitzer S *et al.* (2014) MtDNA segregation in heteroplasmic tissues is common in vivo and modulated by haplotype differences and developmental stage. Cell reports 7, 2031-41.

Chanda R, Ghosh S (2013) The Punjabi Diaspora in the UK: An Overview of Characteristics and Contributions to India, CARIM-India RR 2013/08. Robert Schuman Centre for Advanced Studies, San Domenico di Fiesole (FI): European University Institute.

Deason ML, Salas A, Newman SP, Macaulay VA, St AMEY, Pitsiladis YP (2012) Interdisciplinary approach to the demography of Jamaica. BMC evolutionary biology 12, 24.

Dubut V, Murail P, Pech N, Thionville MD, Cartault F (2009) Inter- and extra-Indian admixture and genetic diversity in reunion island revealed by analysis of mitochondrial DNA. Ann Hum Genet 73, 314-34.

Fendt L, Rock A, Zimmermann B, Bodner M, Thye T, Tschentscher F, Owusu-Dabo E, Gobel TM, Schneider PM, Parson W (2012) MtDNA diversity of Ghana: a forensic and phylogeographic view. Forensic Sci Int Genet 6, 244-9.

Helgason A, Hickey E, Goodacre S, Bosnes V, Stefansson K, Ward R, Sykes B (2001) mtDna and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. American journal of human genetics 68, 723-37.

Johnston IG, Burgstaller JP, Havlicek V, Kolbe T, Rulicke T, Brem G, Poulton J, Jones NS (2015) Stochastic modelling, Bayesian inference, and new in vivo measurements elucidate the debated mtDNA bottleneck mechanism. eLife 4.

Kivisild T, Reidla M, Metspalu E, Rosa A, Brehm A, Pennarun E, Parik J, Geberhiwot T, Usanga E, Villems R (2004) Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears. American journal of human genetics 75, 752-70.

McEvoy B, Richards M, Forster P, Bradley DG (2004) The Longue Duree of genetic ancestry: multiple genetic marker systems and Celtic origins on the Atlantic facade of Europe. American journal of human genetics 75, 693-702.

Owen D (unpublished) African migration to the UK. <https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiWnufU6u_LAhUHow4KHV8RDWEQFggdMAA&url=https%3A%2F%2Fwww2.warwick.ac.uk%2Ffac%2Fsoc%2Fcrer%2Fevents%2Fafrican%2Fconfp_david_owen.ppt&usg=AFQjCNEWDLvQzhy1YcK1Wxh1NKx-77tamQ&sig2=i3v9qAbv5qZxTlenIESOBQ>.

Pharaoh R (2009) Migration, Integration, Cohesion - New Chinese Migrants to London.

Plaza S, Calafell F, Helal A, Bouzerna N, Lefranc G, Bertranpetit J, Comas D (2003) Joining the pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean. Ann Hum Genet 67, 312-28.

Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, Rengo C, Al-Zahery N, Semino O, Santachiara-Benerecetti AS *et al.* (2004) Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. American journal of human genetics 74, 827-45.

Relethford JH, Crawford MH (2013) Genetic drift and the population history of the Irish travellers. American journal of physical anthropology 150, 184-9.

Røyrvik EC (unpublished).

Salas A, Richards M, De la Fe T, Lareu MV, Sobrino B, Sanchez-Diz P, Macaulay V, Carracedo A (2002) The making of the African mtDNA landscape. American journal of human genetics 71, 1082-111.

Saunier JL, Irwin JA, Strouss KM, Ragab H, Sturk KA, Parsons TJ (2009) Mitochondrial control region sequences from an Egyptian population sample. Forensic Sci Int Genet 3, e97-103.

TheChangeInstitute (2009) The Pakistani Muslim Community in England - Understanding Muslim Ethnic Communities. Wetherby.

Xue F, Wang Y, Xu S, Zhang F, Wen B, Wu X, Lu M, Deka R, Qian J, Jin L (2008) A spatial analysis of genetic structure of human populations in China reveals distinct difference between maternal and paternal lineages. European journal of human genetics : EJHG 16, 705-17.