



E-ISSN: 2278-4136
P-ISSN: 2349-8234
JPP 2018; 7(2): 2334-2236
Received: 25-01-2018
Accepted: 26-02-2018

Amit Kaushik
Amity Institute of
Biotechnology, Amity
University Uttar Pradesh,
Noida, Uttar Pradesh, India

Aarti Sharma
Amity Institute of
Biotechnology, Amity
University Uttar Pradesh,
Noida, Uttar Pradesh, India

Sudhir Sharma
Department of Plant Breeding,
College of Agriculture, CCSHAU,
Hisar, Haryana, India

Haplotype: Alleles moving together

Amit Kaushik, Aarti Sharma and Sudhir Sharma

Abstract

Haplotype is a set of single nucleotide polymorphisms (SNPs) on a single chromatid that are statistically associated. Combination of alleles of closely linked loci, those are found in a single chromosome and tend to be inherited together. Haplotype analysis is useful in identifying recombination events. The study of haplotypes within genes, which is also of great current interest, provides the opportunity to discover reliable markers of various phenotypes. Haplotype trees provide methods for examining the phylogeny of individuals based on their haplotypes and also provide methods for understanding molecular (genomic) natural selection. These trees can be created for one species or can be created to represent inter-species haplotypic phylogenies.

Keywords: haplotype, alleles, SNP'S, phylogeny

Introduction

The term haplotype is a contraction of the term "haploid genotype". In genetics, a haplotype (Greek haploos = single) is a combination of alleles at multiple loci that are transmitted together on the same chromosome. Haplotype may refer to as few as two loci or to an entire chromosome depending on the number of recombination events that have occurred between a given set of loci.

In a second meaning, haplotype is a set of single nucleotide polymorphisms (SNPs) on a single chromatid that are statistically associated. It is thought that these associations, and the identification of a few alleles of a haplotype block, can unambiguously identify all other polymorphic sites in its region. Such information is very valuable for investigating the genetics behind common diseases.

The examination of single chromosome sets (haploid sets), as opposed to the usual chromosome pairings (diploid sets), is important because mutations in one copy of a chromosome pair can be masked by normal sequences present on the other copy. Combination of alleles of closely linked loci, those are found in a single chromosome and tend to be inherited together. The linear, ordered arrangement of alleles on a chromosome. Haplotype analysis is useful in identifying recombination events.

Haplotyping involves grouping subjects by haplotypes, or particular patterns of sequential SNPs, found on a single chromosome.

Genomic variation, and thus SNPs, is responsible for diversity in the human species. It follows that since SNPs account for diversity in human genotypes, they can be mapped to account for diversity in phenotypes. An individual's SNPs may serve as signposts for disease genes, haplotypes are believed to be superior for this purpose. The study of haplotypes within genes, which is also of great current interest, provides the opportunity to discover reliable markers of various phenotypes". This relation forms the basis and motivation for the identification and genotyping of SNPS.

Haplotype trees (Haplotype-based phylogeny)

Haplotype trees provide methods for examining the phylogeny of individuals based on their haplotypes and also provide methods for understanding molecular (genomic) natural selection. They are constructed to understand human evolution, historical timelines and to genetically determine genealogy. These trees can be created for one species or can be created to represent inter-species haplotypic phylogenies (Templeton, 2005) [4]. Recall that there are a small number of haplotypes (unique patterns) for chosen location of interest. A population or haplotype group is a set of highly similar haplotypes. Often the haplotype under consideration is a maternally inherited gene or a set of locations on one of the sex chromosomes. It has been shown that members of a population generally share the same haplotype pattern. These trees are often combined with homology-based trees to provide a more reliable portrait of genealogies.

Correspondence
Sudhir Sharma
Department of Plant Breeding,
College of Agriculture, CCSHAU,
Hisar, Haryana, India

Haplotype trees are constructed parsimoniously with unique haplotypes being represented by the nodes of the tree. Haplotypes develop from older ancestral haplotypes. These older haplotypes are believed to be more widespread over the species, and are therefore generally represented by internal nodes, whereas newer haplotypes (more recently emerged patterns) will be represented by leaf nodes. Note that it is integral to select a haplotype that is robust to recombination and mutations.

Family Tree DNA provided the following thumbnail summaries of the different haplogroups:

1. Haplogroup A (M91) is a Y-chromosome haplogroup. Haplogroup A is common among Bushmen, Khoisan, Sudanese, Ethiopians (especially Beta Israel) and Nilotes (Cruciani *et al* 2011) ^[1]. Haplogroup A is localized mainly to Southern Africa with a small to notable presence among a few populations in East Africa. It represents the oldest and most diverse of the human Y-chromosome haplogroups. It is believed to be the haplogroup corresponding to Y-chromosomal Adam.
2. Haplogroup B (M60) is one of the oldest Y-chromosome lineages in humans. Haplogroup B is found exclusively in Africa. This lineage was the first to disperse around Africa. There is current archaeological evidence supporting a major population expansion in Africa approximately 90-130 thousand years ago. It has been proposed that this event may have spread Haplogroup B throughout Africa (Hassan *et al*, 2008) ^[5]. Haplogroup B appears at low frequency all around Africa, but is at its highest frequency in Pygmy populations.
3. Haplogroup C (M130, M216) is found throughout mainland Asia, the south Pacific, and at low frequency in Native American populations. Haplogroup C originated in southern Asia and spread in all directions. This lineage colonized New Guinea, Australia, and north Asia, and currently is found with its highest diversity in populations of India (Zhong *et al*, 2011).
4. Haplogroup C3 (M217, P44) is believed to have originated in southeast or central Asia. This lineage then spread into northern Asia, and then into the America.
5. Haplogroup C5 (M356) Found in the Indian subcontinent.
6. Haplogroup D2 (M55, M57, M179, P12) most likely derived from the D (M174) lineage in Japan, Tibet, the Andaman Islands. It is completely restricted to Japan, and is a very diverse lineage within the aboriginal Japanese and in the Japanese population around Okinawa (She *et al*, 2008).
7. Haplogroup E3a (M2) is an Africa lineage. It is currently hypothesized that this haplogroup dispersed south from northern Africa within the last 3,000 years, by the Bantu agricultural expansion. E3a is also the most common lineage among African Americans (Underhill *et al*, 2001) ^[12].
8. Haplogroup E3b (M215) is believed to have evolved in the Middle East. It expanded into the Mediterranean during the Pleistocene Neolithic expansion. It is currently distributed around the Mediterranean, southern Europe, and in north and east Africa.
9. Haplogroup G (M201) is a branch of Haplogroup F (M89) and is theorized to have originated, according to the latest thinking, in the Near East or Southern Asia, probably in the region that is now northern India, Pakistan, and Afghanistan. The haplogroup began to spread with the Neolithic Agricultural Revolution (Larmuseau *et al*, 2014) ^[10].
10. Haplogroup H (M69) is nearly completely restricted to India, Sri Lanka, and Pakistan (Sengupta *et al*, 2006) ^[11].
11. Haplogroups I, I1 (M253, M307, P30, and P40) and I1a (M21) are nearly completely restricted to northwestern Europe. These would most likely have been common within Viking populations. One lineage of this group extends down into central Europe.
12. Haplogroup I1b (M227) was derived within Viking/Scandinavian populations in northwest Europe and has since spread down into southern Europe where it is present at low frequencies.
13. Haplogroup J (M304, S6, S34, S35) is found at highest frequencies in Middle Eastern and North African populations where it most likely evolved. Middle Eastern traders into Europe, central Asia, India, and Pakistan have carried this marker (Poznik *et al*, 2016) ^[9].
14. Haplogroup J2 (M172) originated in the northern portion of the Fertile Crescent where it later spread throughout central Asia, the Mediterranean, and south into India. As with other populations with Mediterranean ancestry this lineage is found within Jewish populations. The Cohen modal lineage is found in Haplogroup J2.
15. Haplogroup Q (MEH2, M242, P36) is the lineage that links Asia and the Americas. This lineage is found in North and Central Asian populations as well as Native Americans. This lineage is believed to have originated in Central Asia and migrated through the Altai/Baikal region of northern Eurasia into the Americas (Zegura *et al*, 2004) ^[8].
16. Haplogroup Q3 (M3) contains the patrilineal ancestors of almost all of the indigenous peoples of the Americas. Haplogroup Q3 is a descendant of Haplogroup Q (M242). Since the parent Q (M242) appears in both Asia and the Americas, while Q3 (M3) appears only in the Americas, it is believed that the mutation M3 occurred in North America after migration across the Bering Strait had ceased.
17. Haplogroup R1 (M173) Found throughout western Eurasia (Raghavan *et al*, 2014) ^[3].
18. Haplogroup R1a (M17) is believed to have originated in the Eurasian Steppes north of the Black and Caspian Seas. This lineage is believed to have originated in a population of the Kurgan culture, known for the domestication of the horse (approximately 3000 B.C.E.). These people were also believed to be the first speakers of the Indo-European language group (Karafet *et al*, 2014) ^[7]. This lineage is currently found in central and western Asia, India, and in Slavic populations of Eastern Europe.
19. Haplogroup R1b (M343) is the most common haplogroup in European populations. It is believed to have expanded throughout Europe as humans re-colonized after the last glacial maximum 10-12 thousand years ago. This lineage is also the haplogroup containing the Atlantic modal haplotype (HG1).
20. Haplogroup R2 (M124) is characterized by genetic marker M124, and is rarely found outside India, Pakistan, Iran, and southern Central Asia. Haplogroup R2 arose about 25,000 years ago in southern Central Asia, and its members migrated southward as part of the second major

wave of human migration into India. At least 90% of R2 individuals are located in the Indian sub-continent. It is also reported in Caucasian and Central Asian populations.

References

1. Cruciani F, Trombetta B, Massaia A, Destro-Bisol G, Sellitto D, Scozzari R. A revised root for the human Y chromosomal phylogenetic tree: the origin of patrilineal diversity in Africa. *Am. J Hum. Genet.* 2011; 88(6):814-8.
2. Zhong H, Shi H, Qi XB *et al.* Global distribution of Y-chromosome haplogroup C-M130 reveals the prehistoric migration routes of African exodus and early settlement in East Asia. *J Hum. Genet.* 2010; 55(7):428-35.
3. Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature*, 2014; 505(7481):87-91.
4. Templeton AR. Haplotype Trees and Modern Human Origins. *Yearbook of Physical Anthropology.* 2005; 48:33-59.
5. Hassan HY, Underhill PA, Cavalli-Sforza LL, Ibra ME. Y-chromosome variation among Sudanese: Restricted gene flow, concordance with language, geography, and history. *American Journal of Physical Anthropology.* 2008; 137(3):316-323.
6. Shi H, Zhong H, Peng Y, *et al.* Y chromosome evidence of earliest modern human settlement in East Asia and multiple origins of Tibetan and Japanese populations. *BMC Biol.* 2008; 6:45.
7. Karafet T, Fernando M, Herawati S. Improved phylogenetic resolution and rapid diversification of Y-chromosome haplogroup K-M526 in Southeast Asia. *Nature.* 2014; 23:369-373.
8. Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF. High-Resolution SNPs and Microsatellite Haplotypes Point to a Single, Recent Entry of Native American Y Chromosomes into the Americas (PDF). *Molecular Biology and Evolution.* 2004; 21(1):164-75.
9. Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, *et al.* The 1000 Genomes Project Consortium, Bustamante CD, Tyler-Smith C Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences *Nature Genetics.* 2016; 48:593-599.
10. Larmuseau MH, Delorme P, Germain P, Vanderheyden N, Gilissen A, Van Geystelen A *et al.* Genetic genealogy reveals true Y haplogroup of House of Bourbon contradicting recent identification of the presumed remains of two French Kings. *Eur J Hum Genet.* 2014; 22:681-7.
11. Sengupta S, Zhivotovsky LA, King R. Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *Am. J Hum. Genet.* 2006; 78(2):202-21.
12. Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, *et al.* The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Annals of Human Genetics.* 2001; 65(1):43-62.