Realia et naturalia

DISSERTATIONES
BIOLOGICAE
UNIVERSITATIS
TARTUENSIS
247

GEORGI HUDJAŠOV

Maps of mitochondrial DNA, Y-chromosome and tyrosinase variation in Eurasian and Oceanian populations

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
247

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
247

GEORGI HUDJAŠOV

Maps of mitochondrial DNA, Y-chromosome and tyrosinase variation in Eurasian and Oceanian populations

UNIVERSITY OF TARTU
PRESS
1632

Institute of Molecular and Cell Biology, University of Tartu

Dissertation is accepted for the commencement of the degree of Doctor of Philosophy in Molecular and Cell Biology on September 9, 2013, by the Council of the Institute of Molecular and Cell biology, Faculty of Science and Technology, University of Tartu.

Supervisors: Toomas Kivisild,
Professor, Institute of Molecular and Cell Biology,
Faculty of Science and Technology, University of Tartu, Estonia
and
University Lecturer, Division of Biological Anthropology,
University of Cambridge, United Kingdom

Prof. Richard Villems,
Institute of Molecular and Cell Biology,
Faculty of Science and Technology, University of Tartu,
Estonia

Reviewer: Prof. Ants Kurg,
Institute of Molecular and Cell Biology,
Faculty of Science and Technology, University of Tartu, Estonia

Opponent: Prof. Peter de Knijff,
Department of Human Genetics,
Center for Human and Clinical Genetics,
Leiden University Medical Center, The Netherlands

Commencement: Room 105, 23B Riia street, Tartu, on 7th of November 2013, at 12.00.

The publication of this dissertation is granted by the University of Tartu

ISSN 1024–6479
ISBN 978–9949–32–402–6 (print)
ISBN 978–9949–32–403–3 (pdf)

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

I. **Hudjashov G**, Kivisild T, Underhill PA, Endicott P, Sanchez JJ, Lin AA, Shen P, Oefner P, Renfrew C, Villems R and Forster P. 2007. Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. *Proc Natl Acad Sci U S A*. 104: 8726–8730.

II. Soares P, Trejaut JA, Loo JH, Hill C, Mormina M, Lee CL, Chen YM, **Hudjashov G**, Forster P, Macaulay V, Bulbeck D, Oppenheimer S, Lin M and Richards MB. 2008. Climate change and postglacial human dispersals in southeast Asia. *Mol Biol Evol*. 25: 1209–1218.

III. **Hudjashov G**, Villems R and Kivisild T. 2013. Global patterns of diversity and selection in human tyrosinase gene. *PLoS ONE*. 8: e74307.

My contributions to the listed articles are as follows:

Ref. I performed most of the experiments and statistical analysis of the data, wrote the first draft of the paper.

Ref. II performed experiments and statistical analysis of the data.

Ref. III designed the study, performed experiments and statistical analysis of the data, wrote the paper with the contribution of other co-authors.

ABBREVIATIONS

bp	base pair(s)
BP	before the present
CI	confidence interval
hg	haplogroup(s)
ISEA	Island Southeast Asia
kb	thousand (kilo-) base pairs
ky(a)	thousand (kilo-) years (ago)
LGM	Last Glacial Maximum
mtDNA	mitochondrial DNA
NRY	non-recombining region of the Y chromosome
SNP	single nucleotide polymorphism
TMRCA	time to the most recent common ancestor
UV(R)	ultraviolet (radiation)

1. INTRODUCTION

Analysis of genetic variation in modern human populations has become a valuable tool for the study of the prehistory of our species. Recent genetic work has offered new insights into various evolutionary processes, including peopling of the world by anatomically modern humans, but also allowed for better characterization of diverse adaptive processes which took place during the prehistory of living hominins. Thanks to advancements in cost-effective sequencing and high-coverage whole-genome genotyping techniques, an unprecedented amount of population genetic data has been generated during the last decade, covering different human population groups and addressing numerous important evolutionary questions.

Genetic studies on human evolution are tightly interconnected with an array of scientific disciplines, including, among others, anthropology, archaeology and linguistics. For example, the prehistory of Australia and New Guinea has been extensively studied during the last decades. Fossil evidence for the presence of anatomically modern humans in this area has been dated to approximately 50 kya, showing that Greater Australia was probably colonized soon after the African exodus. Meanwhile, the overall knowledge of genetic variation in Australia and New Guinea has been relatively scarce until recently. The first part of this dissertation concentrates on the characterization and classification of haploid mitochondrial DNA (mtDNA) and Y-chromosome lineages among the aboriginal populations of Australia and New Guinea. Complete mitochondrial genome sequencing and Y-chromosomal haplotyping have been utilized in order to place the respective gene pools into a global framework of the corresponding haploid genetic systems.

The second part of the dissertation addresses a different scientific problem. There is little evidence on the evolution of human skin color from classical archaeology and paleoanthropology. Only limited and indirect knowledge can be acquired from the cave paintings and other cultural artifacts. Biological studies of model organisms and albinism have highlighted hundreds of genetic loci, which may also determine normal skin color variation in humans. Nevertheless, only a small fraction of this phenotypic diversity is currently understood. The second part of the current research concentrates on the analysis of the human tyrosinase gene. This gene is a cornerstone of melanin biosynthesis and might play a significant role in shaping natural skin color variation. Questions of natural selection acting upon human tyrosinase will be addressed at both the local and the genome-wide level, and patterns of genetic diversity within the *TYR* locus will be discussed in terms of the evolutionary architecture of human skin pigmentation.

2. LITERATURE OVERVIEW

2.1. Haploid genetic loci in the study of recent human evolution

Although studies of human populations based on complete genome sequences are starting to emerge (Abecasis et al. 2010; Wong et al. 2013), to a large extent our current understanding of human dispersals and genetic variation within and among populations relies on mtDNA and Y-chromosome haploid markers. These loci have some unique features that make them preferable in the study of different evolutionary processes, including the prehistory of anatomically modern humans.

The human mitochondrial genome is a tiny, about 16.6 kb long, circular double stranded DNA molecule (Anderson et al. 1981; Andrews et al. 1999). Compared to the nuclear genome, mtDNA is highly compact, about 93% of its sequence is coding; there are no introns, intergenic sequences are either absent or very small. Mitochondrial DNA is inherited from the population of mitochondria present in the oocyte just prior to fertilization, so only mother will pass her mtDNA to offspring (Giles et al. 1980). Elimination of male mitochondria early in embryogenesis is responsible for the lack of recombination between maternal and paternal mtDNA molecules. There is a marked reduction in the number of mtDNA molecules being transmitted from mother to offspring, the process is known as mitochondrial genetic bottleneck. The size of the bottleneck is predicted to be around 200 molecules in primordial germ cells, which further differentiate into mature oocytes containing 100,000–700,000 copies of mtDNA in humans. Consequently, the combination of random genetic drift and clonal expansion of mtDNA explain the fast segregation of mutant mtDNA molecules between generations (St John et al. 2010 and references therein). The overall rate of mtDNA heteroplasmy has recently been estimated at around 0.2–2% in humans and, although rare examples of non-pathogenic inherited heteroplasmy were previously known (Ivanov et al. 1996), an ultra-deep re-sequencing approach hints that this inherited low variation may nevertheless contribute to the overall mutation load in any one individual (Payne et al. 2013).

Besides the lack of recombination, another specific feature of mtDNA important for evolutionary studies is its relatively high mutation rate. The non-coding control region of mitochondrial DNA evolves approximately two orders of magnitude faster than the nuclear genome. Calibration of its mutation rate yielded $1.6–2.3 \times 10^{-7}$ *vs.* 0.4×10^{-9} substitutions per site per year for the autosomal genome on average, while the mtDNA coding region evolves at lower rates varying from 1.9×10^{-8} to 8.9×10^{-9} between different protein-coding codon positions, rRNA and tRNA coding regions (Soares et al. 2009; Scally and Durbin 2012 and references therein). In absolute terms, the mitochondrial genome acquires on average one synonymous substitution per 7990 years (Loogväli et al. 2009), thus allowing to estimate the age of relatively recent demographic events in the prehistory of anatomically modern humans.

In contrast to the maternally transmitted mitochondrial genome, Y chromosome is male-specific and thus paternally inherited in humans. It is a part of the XY sex-determination system, containing genes that cause testis development. The human Y chromosome is about 60 million base pairs long. While 95% of its sequence does not recombine during meiosis and is therefore called the non-recombining region of Y (NRY) or the male-specific region, X-Y crossing over is a normal and frequent event in the telomeric pseudoautosomal regions (Jobling and Tyler-Smith 2003 and references therein; Skaletsky et al. 2003). In contrast to the faster evolving segments of mtDNA, the rate of evolution of Y chromosome is estimated at around 1.0×10^{-9} substitutions per site per year, which is comparable to the autosomal average (Xue et al. 2009).

The effective population size of both human mitochondrial DNA and Y chromosome is expected to be one-quarter of that of any autosome and one-third of that of the X chromosome. This feature results in lower sequence diversity in comparison to the autosomal genome and makes them much more prone to genetic drift, which involves random changes in the frequency of alleles from one generation to the next, and founder effect (for review see Jobling and Tyler-Smith 2003). In summary, the evolutionary history of human haploid loci, mtDNA and the male-specific region of Y chromosome, is affected only by mutations accumulating with time, but not by the recombination process. Their uniparental mode of inheritance allows precise gender-specific phylogenetic reconstruction and, due to the higher mutation rate of mitochondrial genome, past demographic events in our species can be effectively dated, while their higher susceptibility to genetic drift due to lower effective population size accelerates their differentiation between human populations.

2.2. Out of Africa to Australasia

Phylogenetic analysis of human matrilineal (mitochondrial) and patrilineal (Y-chromosome) lineages, which is the reconstruction of the hierarchical structure of relationship between variants, allows to infer the order and approximate temporal and spatial point of population descent. The pioneer analyses of worldwide human mtDNA (Brown 1980; Cann et al. 1987) lineages, which were later complemented by evolutionary studies of Y chromosome (Hammer 1994; Jobling et al. 1996; Underhill et al. 1997), showed that the human genetic variation is rooted in the African continent, pointing to the African origin of our species. Comparison of human mitochondrial genome variation to that of Neanderthals shows that the root of the human mtDNA tree lies between the African monophyletic cluster of haplotypes (haplogroup, hg) L0 and all other haplogroups (hgs), including haplogroups L1 to L6 (Figure 1) (Behar et al. 2012). The time to the most recent common ancestor (TMRCA) of all human mitochondria was estimated between 160 and 200 kya (Mishmar et al. 2003; Endicott and Ho 2008; Soares et al. 2009). Y-chromosome analyses have generally yielded slightly younger common ancestor ages, ranging from 101–115 to 140 kya (Cruciani et al. 2011; Wei et al. 2013). Lower diversity in Y could be

due to differences in male and female long-term effective population sizes, but possibly may also reflect our still imperfect knowledge of the mutation rates of the two uniparental genetic systems. Recently, an extremely rare and ancient African branch of the Y-chromosome tree has been detected, shifting the age of the most recent common ancestor back to around 330 kya and pointing to either ancient population sub-structure or introgression of archaic Y chromosomes into anatomically modern Africans (Mendez et al. 2013). Alternatively, the large discrepancy with previous estimates could partially be explained by a too low mutation rate used in the calculations of this study (Sayres 2013).

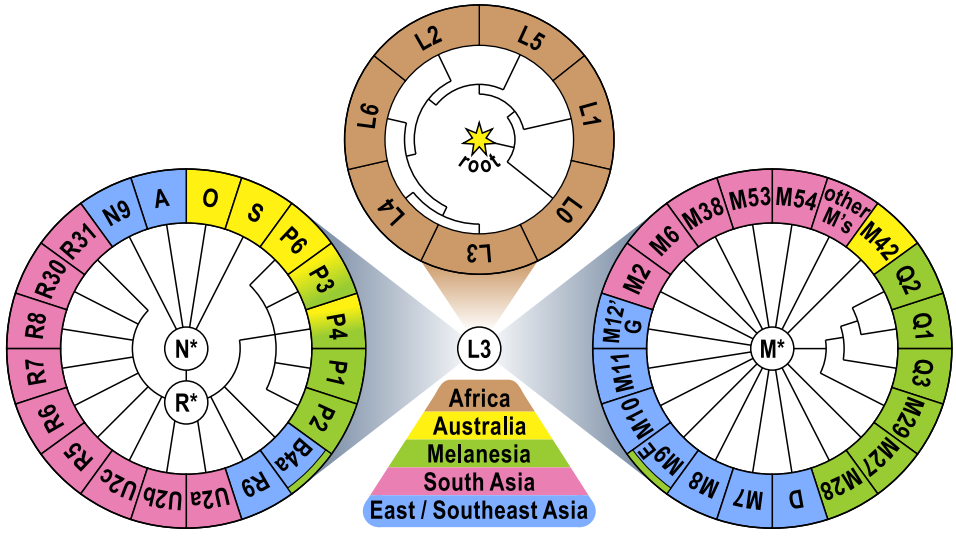

Figure 1. Schematic representation of the global mtDNA phylogenetic tree. Western Eurasian haplogroups are not included. The oldest autochthonous haplogroups from South, East and Southeast Asia, Melanesia and Australia are shown (color coded according to the legend). Hgs E and B4a represent traces of recent gene flow from Asia to Melanesia. Gradient shading of hgs P3 and P4 indicates shared distribution in populations of Australia and Melanesia. The root of the human mtDNA phylogenetic tree lies between the African hgs L0 and L'1–6. All variation outside of Africa is represented by macro-haplogroups M and N, and is rooted in the African hg L3. Note that only a small sub-set of Asian-specific haplogroups is depicted. Greater Australian branches represented only by a single complete mitochondrial genome are omitted. Data from: Ingman et al. (2000), Ingman and Gyllensten (2003), Kong et al. (2003), Palanichamy et al. (2004), Friedlaender et al. (2005), Macaulay et al. (2005), Kivisild et al. (2006), van Holst Pellekaan et al. (2006), Friedlaender et al. (2007), Behar et al. (2008) and Chandrasekar et al. (2009). The latest high-resolution phylogenetic tree of all human maternal lineages can be found elsewhere (van Oven and Kayser 2009).

The oldest African mitochondrial hg L0 lineages coalesces at about 140 to 160 kya (Behar et al. 2008; Soares et al. 2009), while the oldest bifurcation con-

necting both African and non-African individuals occurs in hg L3, which dates back to around 60 to 70 kya (Torroni et al. 2006; Scally and Durbin 2012 and references therein; Soares et al. 2012). Virtually all human mtDNA variation outside of Africa can be divided into two branches, namely macro-haplogroups M and N, the latter including the macro-haplogroup R (Figure 1). This is mirrored in the Y-chromosome tree, where only hgs C, D and F are found in non-Africans, except in likely recent migrants (for review see Underhill and Kivisild 2007). Both mitochondrial M and N trunks of the tree stem from the African hg L3. The age of Eurasian M and N founders is close to the age of their ancestral haplogroup, L3, putting the timeframe for the out of Africa dispersal of anatomically modern humans between 55 and 65 kya (Torroni et al. 2006; Soares et al. 2009; Soares et al. 2012). Similar estimates for human expansion in Eurasia, between 57 and 74 kya, were obtained by complete re-sequencing of Y chromosomes (Wei et al. 2013).

Although the precise location and number of migration waves from Africa to Eurasia are still being debated (Rasmussen et al. 2011; Henn et al. 2012), one model that has been suggested on the grounds of archaeological, paleoclimatic, geographic and genetic evidence is that of a southern route dispersal. According to this model, the dispersal started from the Horn of Africa and lead on to South Arabia, crossing the narrow mouth of the Red Sea, and further on towards South and Southeast Asia along the shoreline of the Indian Ocean, reaching Greater Australia (joint Pleistocene landmass of Australia and northern Melanesia, including New Guinea; also referred to as Sahul) approximately 50 kya (Macaulay et al. 2005; Oppenheimer 2012 and references therein). This dispersal of modern humans was accompanied by replacement with limited gene flow of local archaic human species, such as Neanderthals (Green et al. 2010) and Denisovans (Reich et al. 2010; Reich et al. 2011; Meyer et al. 2012).

Some general geographical patterns of the spread (i.e. phylogeography) of the mitochondrial macro-haplogroups M and N must be emphasized here: macro-haplogroup M and its subdivisions are found primarily in South and East Asia, while macro-haplogroup N is dominant in West Eurasia and Oceania, and to a lesser extent in East Asia (Metspalu et al. 2004; Underhill and Kivisild 2007 for review). Apart from macro-haplogroups M and N, no other hg L3 branches are present outside Africa, suggesting that the earliest Eurasian colonizers already carried basal M and N variants. The N branch quickly gave rise to the third human mitochondrial macro-haplogroup, hg R. During the southern route migration, the differentiation of basal M, N and R lineages into mtDNA branches specific to South, East and Southeast Asia and Oceania took place (Figure 1). On the other hand, only the N branch (and its sub-clade R) gave rise to western Eurasian mtDNA types (Metspalu et al. 2004). Populations from the Indian sub-continent carry diverse ancient, >50 ky old, autochthonous branches of Eurasian macro-haplogroups M, N and R, e.g. M2, M6, R5, U2a'b'c, etc. (Palanichamy et al. 2004; Chandrasekar et al. 2009). The TMRCA of the oldest South Asian haplogroups suggests that this area was settled soon after the African exodus and points to the major role of Indian sub-continent in expan-

sion of anatomically modern humans in Eurasia (Macaulay et al. 2005; Majumder 2010 for review). The topology and more than 50 ky deep coalescence age of human mtDNA haplogroups in India, East and Southeast Asia and Greater Australia, which are related only on the basal level of macro-haplogroups M, N and R, but not nested inside each other, indicates that the overall southern dispersal should have been relatively rapid. This migration carried ancestral M, N and R types along the shores of the Indian Ocean into South and Southeast Asia and, finally, into prehistoric Sahul, followed by independent *in situ* differentiation of local mtDNA branches as can be seen from the distinct geographical distribution of contemporary mtDNA variation outside Africa (Figure 1) (Kong et al. 2003; Palanichamy et al. 2004; Macaulay et al. 2005; Chandrasekar et al. 2009; Peng et al. 2010; Kong et al. 2011).

2.2.1. Settlement of Sahul: archaeological evidence

The time frame of Out-of-Africa migration has been subject to much debate and inferences have been made from fossil, archaeological and genetic evidence. Apart from the fossils from the Skhul and Qafzeh caves in Israel, no other fossil of anatomically modern human outside Africa has been securely dated to >70 kya. Even the few fossils dated to 50–70 kya are often debated and much of the controversy and uncertainty around early dates is due to the limits of the radiocarbon method. One of the oldest anatomically modern human remains in Eurasia have been found in the Callao Cave, Philippines, and dated to approximately 67 kya (Mijares et al. 2010; Oppenheimer 2012 and references therein), while Greater Australia, which is the easternmost end of the Out-of-Africa migration (the Americas and Remote Oceania were settled much later), has a number of sites with dates up to 49 kya (Bowler et al. 2003; O'Connell and Allen 2004; Summerhayes et al. 2010). The lack of ancient archaeological sites on the route of human migration to Southeast Asia and Sahul, along the coast of the Indian Ocean, is not surprising, bearing in mind that much of the Late Pleistocene evidence from the coastal area of South and Southeast Asia must now be submerged beneath the high sea levels that have persisted since the end of the last glacial period approximately 12.5 kya (Stringer 2000).

Pleistocene was a period of fluctuating sea level. Starting from the beginning of the last glacial cycle, some 120 kya, the sea level fell gradually to an extremely low stand during the Last Glacial Maximum (LGM) about 18 kya. The modern landmasses of Australia, Tasmania, New Guinea, and intervening islands were connected by a land bridge forming the prehistoric Sahul continent (Figure 2). Ice melting and rising sea level after the end of the Last Glacial Maximum have led to the inundation of lowland territories and the separation of islands from mainland Australia: Tasmania became separated by the Bass Strait approximately 10 kya and New Guinea was connected with the continent until 8 kya, when the Torres Strait started to emerge (Mulvaney and Kamminga 1999). The oldest archaeological sites, including Lake Mungo and the Devil's Lair (47–42 ky) in southeastern and southwestern Australia, respectively, Ivane

Valley (49–43 ky) in eastern New Guinea, and several sites in the Bismarck Archipelago near the northeastern tip of Sahul (up to 45 ky) are all localized in distant parts of the prehistoric continent (Figure 2), implying that both present-day Australia and New Guinea were rapidly occupied by anatomically modern humans at about the same time (Mulvaney and Kamminga 1999; O'Connell and Allen 2004; Summerhayes et al. 2010). Therefore, it is clear that approximately 50 kya was the latest possible time for human arrival in Greater Australia (Hiscock 2008).

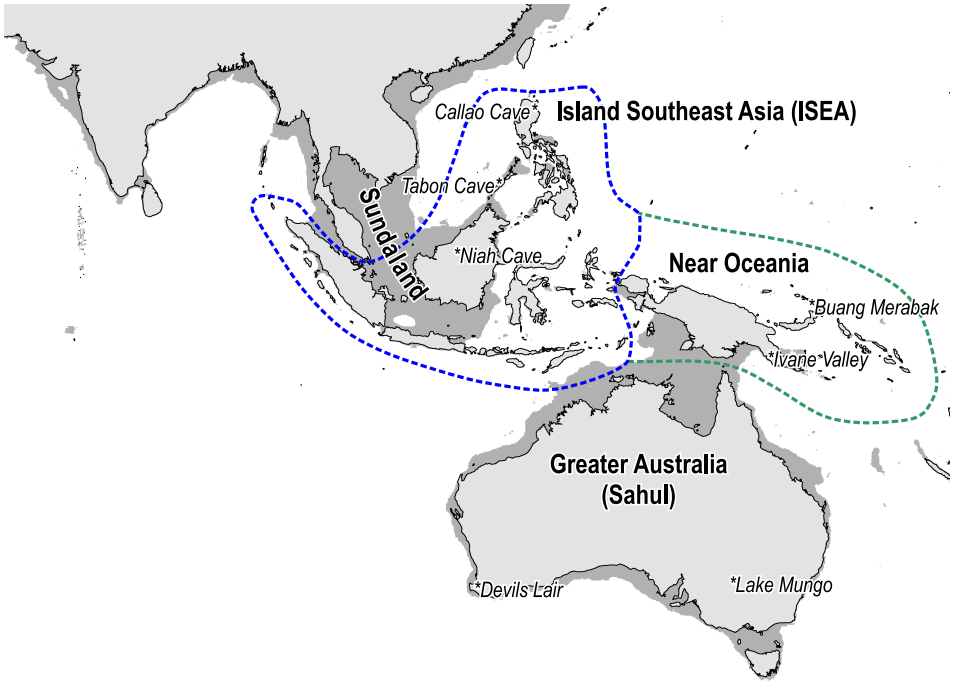

Figure 2. Map of Southeast Asia and Australia, showing both the modern coastline (light gray) and the 120-m depth contour below the current sea level (dark gray), and indicating the extent of Late Pleistocene land masses. Two prehistoric (sub-)continents are shown: (a) Sundaland – joint land mass of mainland Southeast Asia, Sumatra, Borneo, Java, Bali and surrounding smaller islands, and (b) Sahul or Greater Australia – joint land mass of Australia, New Guinea, Tasmania and intervening islands. The border between Sundaland and Sahul now passes along Wallacea, a biogeographical designation for a group of Indonesian islands separated by deep water from the Asian and Australian continental shelves. Also shown are Near Oceania (northern Melanesia, including New Guinea, the Bismarck Archipelago and the Solomon Islands) and the region of Island Southeast Asia (ISEA). Asterisks indicate archaeological sites with the oldest dates of regional presence of anatomically modern human remains (adapted from Oppenheimer 2012).

2.2.2. Settlement of Sahul: a haploid perspective

Aboriginal Australians are currently one of the most underrepresented populations in the genetic research: for example, only 38 complete mtDNA sequences are available for Aboriginal Australians out of at least 16,810 complete mitochondrial genomes listed in Build 15 of the PhyloTree database on 30th September, 2012 (van Oven and Kayser 2009). The main limiting factor for their study is a very restricted access of the scientific community to endogenous Australian samples due to the distrust and resistance to genetic studies of many local people. This is caused by various problems, including poor research practices of the past, failure to present the benefits of genetic research to its participants, discomfort caused by publication of intimate ancestry data, fears that genetic research might produce results that erode identity, but also the reluctance of scientists for lengthy and tedious negotiation and engagement processes required by strict local regulatory bodies (van Holst Pellekaan 2012). Many already published sample sets are shared between successive studies and were collected from a limited number of geographical locations long before the advancement of large-scale DNA sequencing techniques.

2.2.2.1. mtDNA variation in Australia and New Guinea

Although a number of authors have analyzed the partial control region sequence variation (Stoneking et al. 1990; Vigilant et al. 1991; Lum et al. 1994; Sykes et al. 1995; Betty et al. 1996; van Holst Pellekaan et al. 1998; Redd and Stoneking 1999; Lum and Cann 2000; Forster et al. 2001; Huoponen et al. 2001; Tommaseo-Ponzetta et al. 2002), our knowledge about complete mtDNA and NRY sequence phylogeny in Aboriginal Australia is still very limited (Ingman et al. 2000; Kayser et al. 2001; Underhill et al. 2001; Redd et al. 2002; Ingman and Gyllensten 2003; Kivisild et al. 2006; van Holst Pellekaan et al. 2006; Friedlaender et al. 2007; Taylor et al. 2012), and genome-wide data is restricted to only a few studies (McEvoy et al. 2010; Rasmussen et al. 2011; Pugach et al. 2013). Taking only matrilineal lineages into consideration, the resolution provided by control region sequencing alone is phylogenetically less reliable than that of complete mtDNA analysis (Ingman et al. 2000). This is caused by a higher frequency of unresolved homoplasy (convergent evolution), which could sometimes be misinterpreted as shared ancestry, and recurrent mutations due to the higher control region mutation rate (Soares et al. 2009). The same problems also occur in the genealogy of complete mtDNA sequences, but to a far lesser extent, and complete mitochondrial genome analysis has become a standard tool since the advancement of more cost-effective sequencing techniques (Torroni et al. 2006). The main topic that will be addressed here is high-resolution complete mtDNA sequence phylogeny of endogenous matrilineal lineages in the ancient Sahul area and, in particular, in Australia, and its links to Eurasian mtDNA variation. As noted above, there is very limited matrilineal and patrilineal genetic information available from different Aboriginal Australian popu-

lations. The majority of the data were generated using outdated partial control region sequencing sometimes complemented by the analysis of a few coding region SNPs. This information does not always allow comprehensive assignment of different mtDNA types to haplogroups or may even lead to cladistic misclassifications of identified haplotypes.

The landscape of mtDNA haplogroup frequencies in the extant populations of prehistoric Sahul is characterized by the presence of deep autochthonous lineages that are not found elsewhere in the world and a clear distinction of the haplogroup profiles of Aboriginal Australian and New Guinean populations. These autochthonous haplogroups of Sahul derive, however, from the same two mtDNA macro-haplogroups (M and N) characteristic to virtually all anatomically modern humans outside of Africa. High nucleotide diversity, compatible to that of the Asian population, points to the great antiquity of human groups inhabiting different areas of Greater Australia (Ingman and Gyllensten 2003). Data indicates that there is only limited lineage sharing between populations from Australia and northern Melanesia and that endogenous people from highland New Guinea and Australia are more closely related to each other than to any other Asian population. On the contrary, populations from coastal New Guinea have higher genetic affinity with populations from East and Southeast Asia and Polynesia (Ingman and Gyllensten 2003; Friedlaender et al. 2007). This Asian legacy can likely be traced back to the recent mid-Holocene migration and represents a common genetic pool of mitochondrial hg B4a subbranches: hgs B4a1a1a (the so-called Polynesian motif) and B4a1a1 (Figure 1). These lineages are found in Remote Oceania, coastal New Guinea, Island Southeast Asia (ISEA) and Taiwan, but are virtually missing in Aboriginal Australians and rare in New Guinea highlanders (Friedlaender et al. 2007; Kayser 2010 and references therein; Soares et al. 2011). Hg B4a sub-branches were probably introduced to Near Oceania during the Austronesian language expansion, which, according to some models, started from the region of southern China and Taiwan about 5,000 to 6,000 years ago and resulted in the settlement of Remote Oceania and Polynesia (Merriwether et al. 2005; Trejaut et al. 2005). The 9-bp deletion at positions 8281–8289 that is characteristic to all hg B lineages was found at very low frequency among Australian Aborigines. Nevertheless, the lack of additional downstream hg B4a defining mutations points to its convergent evolutionary history in Australia, which is further supported by the high incidence of this 9-bp deletion in worldwide human mtDNA phylogeny (Betty et al. 1996; van Oven and Kayser 2009).

Additionally, another maternal haplogroup that is occasionally found in Near Oceania, but not in Australia, is hg E, a sub-branch of M9 (Figure 1) (Merriwether et al. 2005; Friedlaender et al. 2007; Kayser 2010 and references therein). It has likely reached northern Melanesia during the series of dispersals and expansions of coast-dwelling human populations that began in eastern Sundaland/northwest Wallacea (islands between Sunda and Sahul, including Sulawesi) around the peak of the last glaciation. The expansion was most likely triggered by the rising sea level after the end of Last Glacial Maximum and

entered Near Oceania during the Holocene, probably in association with the Austronesian impact (Friedlaender et al. 2007; Hill et al. 2007).

The most frequent autochthonous maternal lineages in Greater Australia belong to hg Q, a sub-cluster of macro-haplogroup M, and hg P, a sub-cluster of macro-haplogroup R (Figure 1). Hgs P and Q were originally defined using restriction fragment length polymorphisms (Forster et al. 2001), and their internal phylogeny has later been updated by several complete mtDNA analyses (Friedlaender et al. 2005; Merriwether et al. 2005; van Holst Pellekaan et al. 2006; Friedlaender et al. 2007; Corser et al. 2012). Both lineages are rare in populations living west of New Guinea, indicating their Greater Australian homeland. Hg Q shares a single mutation with another Melanesian-specific branch, M29, therefore potentially forming an ancient M29'Q clade (Figure 1), although the possibility of convergent evolution cannon be ruled out (Merriwether et al. 2005). Hgs P and Q are common in Near Oceania (Friedlaender et al. 2005; Friedlaender et al. 2007), and found at lower frequencies in the ISEA region (Hill et al. 2007; Tabbada et al. 2010), thus possibly reflecting shared ancestry between aboriginal populations of Near Oceania and insular Southeast Asia (Mona et al. 2009). Hg Q is characterized by the presence of three sister-clades, Q1–3, whereas hg P is more diverse and includes eight different sub-branches, P1–8, among the populations of prehistoric Sahul. The oldest sub-clades of P and Q (P1, P3 and Q1), along with the Australian-specific M42, are all dated by various methods to approximately 50 kya and earlier, i.e. to the time of the initial settlement of prehistoric Sahul by anatomically modern humans (Friedlaender et al. 2005; van Holst Pellekaan et al. 2006; Friedlaender et al. 2007). Only two sub-lineages, except those on the level of macro-haplo-groups M and N and hg P, namely P3 and P4, are shared between Aboriginal Australian and northern Melanesian populations. Both branches are very deep, suggesting an ancient connection between the two regions and almost complete isolation since then (Ingman and Gyllensten 2003; Friedlaender et al. 2005; van Holst Pellekaan et al. 2006; Friedlaender et al. 2007). While the majority of New Guinean maternal genomes can be classified to hgs P1–2, Q, E and B4a, the Aboriginal Australian mitochondrial landscape carries unique auto-chthonous branches not found elsewhere, including hgs M42, S and O (sub-groups of macro-haplogroup N), and P5–8 (Figure 1) (Palanichamy et al. 2004; Friedlaender et al. 2005; Kivisild et al. 2006; van Holst Pellekaan et al. 2006). Three other deep ancient macro-haplogroup M lineages, M27, M28 and M29, are widespread in Northern Island Melanesia (Bismarck Archipelago and Solomon Islands), but virtually absent in New Guinea and completely lacking in Australia (Merriwether et al. 2005).

Although there are a few dozens of complete mtDNA genomes available from Australia, several problems complicate matters. Firstly, complete mtDNA variation in Aboriginal Australia is, at least partially, not yet established according to the common rules of human mitochondrial phylogeny (e.g. haplogroup names like AuA, AuC, etc. are used instead of S, P4, etc.), and/or the neighbor-joining distance-based tree reconstruction algorithm is sometimes

used instead of the more informative median-joining character state approach (Ingman et al. 2000; Ingman and Gyllensten 2003; van Holst Pellekaan et al. 2006). Secondly, none of the published studies has explicitly analyzed all currently available data, therefore making it difficult to readily incorporate Aboriginal Australian complete mtDNA genomes into the worldwide phylogeny of human matrilineal variation.

2.2.2.2. Y-chromosome variation in Australia and New Guinea

Although the resolution of the NRY tree branches specific to Australia and New Guinea is less developed than that of mitochondrial DNA, paternal variation in prehistoric Sahul and its relation to other Eurasian populations largely supports the results of matrilineal analysis, with some differences. The overall Y-chromosome heterogeneity there appears to be smaller than that of mtDNA: Aboriginal Australians are represented by two main NRY clades: K-M9 and M-M130 (for review see Kayser 2010), with hg C-M130 bearing the Australian-specific DYS390.1 deletion (C-DYS390.1del) reaching up to 69% of the total variation (Figure 3) (Kayser et al. 2001). The spread of hg C-DYS390.1del among different Aboriginal populations across the Australian continent indicates the antiquity of this NRY clade (Forster et al. 1998; Kayser et al. 2001; Redd et al. 2002; Taylor et al. 2012). Contrary to the maternal prehistory of the Aboriginal Australians, a very strong recent introgression of European Y chromosomes was detected in some Aboriginal sample sets, also reflected by whole-genome data (Underhill et al. 2001; McEvoy et al. 2010). For example, 71% out of 757 male samples in the Aboriginal forensic database were assigned to European-specific haplogroups (Taylor et al. 2012). Due to the lower resolution, a considerable amount of male samples can only be classified to the macro-haplogroup K-M9 level, which is frequent outside of Africa (Kayser et al. 2001; Underhill and Kivisild 2007; Taylor et al. 2012). Virtually no Y-chromosome lineages shared between the New Guinean and Aboriginal Australian NRY pools were detected, except those on the level of macro-haplogroups K-M9 and C-M130, the minute frequencies of eastern Indonesian/Melanesian-(C2-M38) and Asian-specific (O-M175) haplogroups likely represent historic admixture in the 20th and 21st centuries (Taylor et al. 2012). This contrasts the distribution of mitochondrial hgs P3 and P4, which are present both in Australia and New Guinea (for review see Kayser 2010).

Although the occurrence of different paternal lineages varies substantially between populations of New Guinea, the following haplogroups can be noted: C2-M38, including C2a-M208, M1-M4, S-M230, and sub-branches of hg O-M175 (Figure 3) (Karafet et al. 2010; Kayser 2010). Hg C2a-M208 has northern Melanesian ancestry; C2-M38, M1-M4 and S-M230 were also detected in eastern Indonesia, pointing to either a more westward origin or gene flow from Near Oceania towards eastern Indonesia, probably the same process that might have introduced the Melanesian-specific mitochondrial hg Q to the ISEA region. In contrast with C, M and S, hg O3a-P201 was likely introduced from East Asia

during the Austronesian language expansion along with the pre-Polynesian B4a1a1 motif of mitochondrial hg B4a, while O1a-M119* Y chromosomes represent pre-Austronesian gene flow, the process probably responsible for the introduction of mtDNA hg E to Near Oceania (Hill et al. 2007; Karafet et al. 2010; Soares et al. 2011). Finally, some New Guinean Y chromosomes cannot be resolved beyond the macro-haplogroup K-M9 level. Leaving aside the K-M9* chromosomes, which most likely form a heterogeneous paraphyletic group of unidentified haplogroups, the only haplogroup common in both regions is C-M130, represented by the sister-clades C-DYS390.1del and C2-M38 (including C2a-M208) in Australia and New Guinea, respectively. Hg C-M130 most likely originated in Asia after the African exodus and stands out as a very ancient link between the two Y-chromosome pools of prehistoric Sahul (Karafet et al. 2008; Kayser 2010 for review).

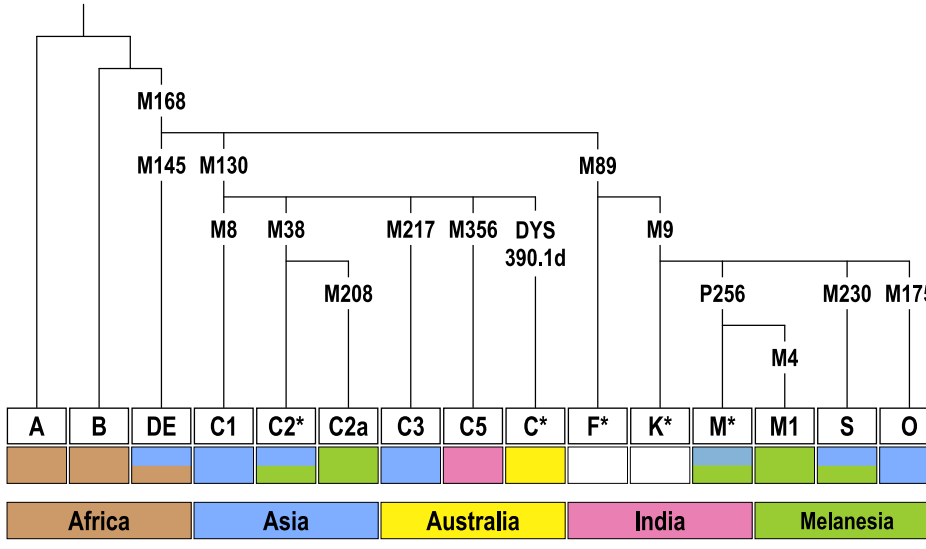

Figure 3. Simplified human Y-chromosome tree. Haplogroups specific to Aboriginal Australia and Melanesia are shown in details, also shown hg C5-M356 found among populations of South Asia. Haplogroup geographical origin is color-coded according to the legend. Modified from Karafet et al. (2008).

The overall picture that emerges from the analysis of human haploid loci is consistent. Both regions of Greater Australia possess ancient autochthonous maternal (Figure 1), paternal (Figure 3) and autosomal variation, supporting the very early occupation dates by anatomically modern humans (Friedlaender et al. 2005; van Holst Pellekaan et al. 2006; McEvoy et al. 2010; Summerhayes et al. 2010; Rasmussen et al. 2011). Aboriginal Australians show a high level of genetic differentiation not only from East and Southeast Asia and Eurasia in general, but likewise from neighboring New Guinea, offering further support

for a rapid southern route dispersal of modern humans to Southeast Asia and Sahul (Ingman and Gyllensten 2003; Friedlaender et al. 2005; Macaulay et al. 2005; Merriwether et al. 2005; Friedlaender et al. 2007; Jinam et al. 2012). Only one maternal haplogroup is shared between Aboriginal Australia and New Guinea, namely hg P (Friedlaender et al. 2005; Friedlaender et al. 2007). This mostly non-overlapping pattern of haplogroup distribution in two geographic regions suggests that the link between modern Aboriginal Australians and New Guineans is very ancient – a conclusion supported independently by the spread of different Y chromosomes in prehistoric Sahul (Kayser et al. 2001; Friedlaender et al. 2005). The distribution and age of the shared haplogroup P imply that both areas were occupied initially by a single migration wave which most likely already carried the derived hg P allele. This founder group, which also included other ancestral M, N and R types, has further split and given rise to the proto-Australian (which inherited all three macro-haplogroup M, N and R (including P) lineages) and proto-Near Oceanian mtDNA pools (inherited only M and R (including P) lineages) (Figure 1). Alternatively, available genetic data is as yet limited to exclude the scenario of multiple independent migration waves that started from closely related source populations at about the same time, but brought separate mtDNA packages to different areas of prehistoric Sahul (Ingman and Gyllensten 2003; Friedlaender et al. 2005; van Holst Pellekaan et al. 2006). Approximately 50 ky old dates for the oldest autochthonous mtDNA haplogroups in Greater Australia do not contradict the age of the first anatomically human remains in this area (O'Connell and Allen 2004; Friedlaender et al. 2005; van Holst Pellekaan et al. 2006; Friedlaender et al. 2007; Summerhayes et al. 2010). Furthermore, the distribution of mitochondrial and Y-chromosome motifs points to the considerable isolation of proto-Australians since the initial split from the proto-Near Oceanian population despite the land bridge connection that persisted until 8 kya. This idea is further supported by the analysis of autosomal loci (Roberts-Thomson et al. 1996; McEvoy et al. 2010) and the phylogeography of a human bacterial parasite, *Helicobacter pylori*, in populations of prehistoric Sahul (Moodley et al. 2009), and prevails in some recent studies by Australian archaeologists (Mulvaney and Kamminga 1999).

2.2.3. Evidence for pre-historic gene flows

The evolutionary history of Aboriginal Australians and New Guineans diverges greatly after the initial colonization of prehistoric Sahul. There are traces of relatively recent admixture between the original inhabitants of Near Oceania, New Guinea in particular, and populations originating from East Asia. The evidence derives from mtDNA, NRY and autosomal analyses, and is supported by the distribution of different language families in this region (Kayser 2010 and references therein). While the New Guinean mainland is dominated by endogenous Papuan-speakers (sometimes referred to as non-Austronesians) representing a very rich linguistic diversity of about 850 languages, the majority of the islanders around New Guinea as well as the inhabitants of the northeast and

southeast coast of New Guinea speak Austronesian languages, which are thought to have arrived from East Asia not earlier than 3,500 years BP (Specht and Gosden 1997; Paul et al. 2013). The Austronesian language expansion is assumed to have started about 5,500 years ago from Taiwan and distributed to Island Southeast Asia, and further to northern Melanesia. These people then continued to expand eastwards into the Pacific. The migration introduced specific mtDNA and NRY lineages to New Guinea and Near Oceania in general, but not to Australia. These included the mtDNA hg B4a1a1, from which most likely *in situ* emerged the so-called Polynesian motif, and the Y-chromosome hg O3a-P201 (Friedlaender et al. 2007; Kayser et al. 2008; Karafet et al. 2010; Soares et al. 2011). Although there is an ongoing debate concerning the demographic *vs.* cultural impact of the Austronesian migration, and the spatial and temporal details of its origin, including which haploid markers are associated with this mid-Holocene migration and which could be ascribed to the early-Holocene/late-Pleistocene gene flow from (insular-)Southeast Asia (Trejaut et al. 2005; Hill et al. 2007; Tabbada et al. 2010; Soares et al. 2011; Jinam et al. 2012; Oskarsson et al. 2012), it is clear that these lineages are not endogenous to the populations of Near Oceania and were introduced there recently. Interestingly, the proportion of Asian-specific types differs significantly between the paternal and maternal gene pools: the average frequency of Asian mtDNA haplogroups varies from 42 to 58% between non-Austronesian and Austronesian speakers of Near Oceania, whereas Asian NRY lineages comprise only between 16 and 2%, respectively, suggesting a female-biased gene flow from incoming Austronesians to local Papuan-speaking populations (Kayser et al. 2008). On the other hand, the abundance of autochthonous Melanesian paternal lineages in Austronesian-speaking groups can be explained by the matrilocal nature of these communities, where non-Austronesian men, rather than women, have moved to Austronesian villages, thereby diluting the original Asian-specific NRY gene pool (Trejaut et al. 2005; Kayser et al. 2008; Kayser 2010).

While Austronesian language expansion has received much attention, there is another possible evidence of post-LGM Southeast Asian gene flow stemming from the distribution of mitochondrial hg E, a subset of Asian-specific M9, and the O1a-M119* Y chromosomes (Hill et al. 2007; Karafet et al. 2010). Hg E is the most common lineage entirely endogenous to ISEA with an average frequency of about 14%. Analysis of mtDNA control region indicates that hg E most likely evolved within the population of the eastern coastline of Sundaland approximately 25 kya and has thereafter experienced serial expansion northwards to Taiwan and, to a lesser extent, eastwards into Near Oceania (Friedlaender et al. 2007; Hill et al. 2007).

Because of a substantially lower amount of genetic data available from Aboriginal Australians, the possibility for gene flow into this population is heavily debated (Brown 2013). In 1870, the outstanding English naturalist Thomas Huxley noted: *"The only people out of Australia who present the chief characteristics of the Australians in a well-marked form are the so-called hill-tribes who inhabit the interior of the Dekhan, in Hindostan"* (Huxley 1870).

This observation has received much attention since. Archaeological data indicates an intensification of the density and complexity of different stone tools in Australia during the mid-Holocene period and the emergence of microlithic technology around 4,000 to 5,000 years ago (Mulvaney and Kamminga 1999 and references therein). The first Dingoes (*Canis lupus dingo*) similarly appear at about the same time (3,500–4,000 years ago) (Gollan 1985) and, along with new stone tool types, were proposed to have been introduced from India by a single package (Glover and Presland 1985; Mulvaney and Kamminga 1999 and references therein). However, later studies have suggested that there is *"no reason to infer that any of the innovations in stone technology derive from overseas"* (Mulvaney and Kamminga 1999). The dating of the oldest backed stone tools has recently shifted to approximately 15,500 years BP, supporting the idea for the independent appearance of the Dingo and microlithic technology in Australia (Hiscock 2008 and references therein). Genetic studies also indicate that the ancestry of the Dingo and its closest relative, the New Guinean singing dog, can be traced back to South China, rather than to the Indian sub-continent. Estimates of the arrival time of the Dingo to Australia from genetic data have a wide range of 4,600 to 18,300 years BP (Savolainen et al. 2004; Ardalan et al. 2012; Oskarsson et al. 2012). Therefore, support for the proposed Indian-Australian mid-Holocene connection remains elusive from the latest archaeological and paleozoological evidence.

Initially, closer affinity of Aboriginal Australians to South Indians rather than to New Guineans was inferred from partial control region sequences of the mitochondrial genome (Redd and Stoneking 1999). However, as already noted above, analysis of control region data alone may be misleading. Indeed, no evidence from complete mtDNA analyses confirmed this early speculation (Ingman and Gyllensten 2003; van Holst Pellekaan et al. 2006). Patrilineal microsatellite analysis similarly suggested that Australian C-M130* Y chromosomes share a mid-Holocene ancestry with populations from Hindustan (Redd et al. 2002). However, later studies confirmed that the majority of the Indian C-M130* chromosomes harbor the ancient M356 derived allele (hg C5), while Aboriginal Australians are characterized by the presence of C-DYS390.1del chromosomes, therefore undermining claims of a recent common ancestry with India (Kayser et al. 2001; Sengupta et al. 2006).

The limited number of whole-genome studies available for the Aboriginal Australian population report conflicting results. McEvoy et al. (2010) found no evidence for any gene flow from outside of Australia, including South India, except the recent admixture with European settlers. The same study also provided support for a shared ancestry of populations from Australia and Melanesia and pointed to the considerable isolation of Australia since the initial Late Pleistocene settlement until the first Europeans arrived in 1788, as confirmed by the single complete genome sequence of an ancient Aboriginal Australian (Rasmussen et al. 2011). On the contrary, traces of an Indian-Australian mid-Holocene connection were detected in a recent genotyping-based study using, among others, structure-like analysis (Pugach et al. 2013). However, the results

of this approach cannot be readily interpreted as evidence of recent admixture between two populations and might as well reflect deep shared proto-Eurasian ancestry (Yunusbayev et al. 2012). Strong marker ascertainment bias, which is inevitably introduced during the design of genome-wide genotyping arrays, should also be taken into the account when studying diverse populations such as the endogenous inhabitants of prehistoric Sahul. It could distort both measures of human diversity and the conclusions drawn from them (Albrechtsen et al. 2010). Therefore, further comprehensive studies of genetic variation, preferentially based on whole genome re-sequencing, including autosomes, Y chromosomes and mtDNA, are needed to better understand the genetic prehistory of Aboriginal Australians and, in particular, to resolve their potentially hitherto hidden connections with other Eurasian populations, including those inhabiting South Asia.

2.3. Out-of-Africa: adaptation to new environments

During their migration out of Africa, anatomically modern humans expanded to vast areas and encountered a diverse range of new environmental conditions, which have triggered a number of novel genetic adaptations. Multiple examples of local adaptive processes are known, including pathogen-driven selection (Fumagalli et al. 2011), adaptation to high altitude (Simonson et al. 2010), climate (Hancock et al. 2008) and diet (Hancock et al. 2010), selection for short stature (Jarvis et al. 2012; Migliano et al. in press), and adaptation to low ultraviolet environment (for review see Sturm and Duffy 2012). One of the most striking differences between modern human populations is skin color. Approximately 88% of the phenotypic variation of this trait lies between different geographical regions in comparison to only 10–15% for craniometrical and genetic loci – a very uncommon pattern, assumed to have been shaped by natural selection (Relethford 2002). The color of our skin is mainly determined by the mix of carotenoids, oxy- and deoxyhemoglobin (erythema), and, most importantly, melanin content. While erythema varies largely during inflammation, thermoregulation and other physiological processes, it is the concentration of different melanin molecules that is responsible for the light to dark color component of human skin. Two types of melanin exist, dark photoprotective pheomelanin and light eumelanin. Both melanins arise from a common metabolic pathway and are synthesized by melanocytes in specialized organelles, melanosomes. Melanosomes are further transferred to keratinocytes residing in the upper layer of the epidermis and thus also determine the color of hair. Differences in skin color do not result from the number of melanocytes (it is equal among different skin types), but from other factors, including the size, number and shape of melanosomes and their aggregation into larger clusters, as well as from the intracellular distribution of melanosomes, melanogenic activity and the rate of melanosome degradation in keratinocytes. For example, light

skin is enriched in light-brown 5,6-dihydroxyindole-2-carboxylic acid (DHICA)-eumelanin and yellow to reddish pheomelanins packed into smaller and densely aggregated melanosomes; on the contrary, dark-brown 5,6-dihydroxyindole (DHI)-eumelanin prevails in the large and non-aggregated melanosomes of heavily pigmented individuals (Jablonski 2004; Parra 2007 and references therein). The concentration of eumelanin, in comparison to pheomelanin, is generally higher in all skin types, while net melanin content varies largely: heavily pigmented black skin has an approximately six-fold higher total melanin load than the lightly pigmented type (Ito and Wakamatsu 2003). Furthermore, melanins differ in their physico-chemical properties. For example, DHI-eumelanin, which is prevalent in dark skin, has the highest ultraviolet (UV) photoabsorption, followed by light-brown DHICA-eumelanin, while pheomelanins have insignificant UV-protecting capacity. The optical properties of a mixture of different melanins are very complex and include the ability to absorb, scatter and reflect light at different wavelengths, including UVA and UVB, therefore acting as a physical barrier to DNA damaging and carcinogenic UV radiation (UVR) (Ortonne 2002; Brenner and Hearing 2008).

The correlation between levels of skin pigmentation and UV radiation is extremely straightforward, pointing to a likely causative connection between the two variables (Chaplin 2004). Several evolutionary drivers have been put forward to explain this association and it is most likely that a combination of various factors has influenced the pigmentation of our skin (for review see: Steindal and Moan 2008; Juzeniene et al. 2009). While Charles Darwin, in his book "The Descent of Man, and Selection in Relation to Sex", ascribed the differences in skin color among humans to sexual selection (Darwin 1871), more than a century later there is as yet no compelling evidence to support this hypothesis, neither sufficient evidence to completely exclude it. The two most renowned hypotheses involving natural selection include: (a) selection for vitamin D_3 biosynthesis which is relevant for explaining the evolution of lighter skin color at higher geographical latitudes with low UVR levels (Loomis 1967), and (b) selection against folate (vitamin B_9) photolysis, and therefore for darker skin color, at lower geographical latitudes, e.g. (near-)equatorial areas with high UVR levels (Branda and Eaton 1978). Vitamin D_3 is a fat-soluble molecule responsible for the regulation of bone health, calcium homeostasis and other physiological processes. It is photochemically produced from 7-dehydrocholesterol in the skin under reaction with UVB. The deficiency of this vitamin is associated with rickets, but likewise involved in cancer, autoimmune diseases, hypertension, and infectious diseases (Holick and Chen 2008; Yuen and Jablonski 2010), while folate is essential for the synthesis and repair of nucleic acids, DNA methylation, and metabolism of amino acids, and its deficiency could lead, among others, to potentially fatal birth defects such as neural tube defects, pregnancy complications and male infertility (National Institutes of Health 2013). The metabolism of both molecules is affected by the photoprotective properties of human skin, which must have been under strong influence of natural selection. Therefore, the pattern of skin color variation among areas

of different UVR load (and geographical latitudes) can be explained as a balance between skin dark enough to protect from folate degradation, but light enough to permit sufficient vitamin D_3 biosynthesis (Jablonski and Chaplin 2010). This is generally supported by insufficient vitamin D_3 levels among heavily pigmented individuals living in UV-poor areas: for example, African Americans have a more than two times lower 25-hydroxyvitamin D concentration than African populations living in their ancestral UV-rich environment (Durazo-Arvizu et al. 2013). However, there are ongoing debates concerning the importance of both vitamin D_3 and folate concentration as evolutionary drivers for human skin (de-)pigmentation. Incidences of rickets in heavily pigmented individuals living at high latitudes are sometimes contradictory and the European archaeological record does not indicate a prevalence of rickets and osteomalacia before the Industrial Revolution and urbanization. The latter suggests that changes in lifestyle and associated environment during the last few hundred years might be causative for this disease (Aoki 2002 and references therein). The folate hypothesis, which was originally proposed based on an *ex vivo* study of folate photodegradation in human plasma, similarly shows inconsistent support from the latest *in vivo* analyses (Juzeniene et al. 2010 and references therein).

Interestingly, there are hypotheses associating skin lightening in northern Europe with the rise and subsequent spread of agriculture starting from about 10–12 kya (Cavalli-Sforza et al. 1994; Juzeniene et al. 2009 and references therein; Khan and Khan 2010). Terrestrial game and fish are rich in vitamin D, while cultivated grain and livestock are a very poor supply of this micronutrient. Therefore, dramatic changes in diet and increased incidence of vitamin D deficiency may have led to additional selection upon skin depigmentation. In addition to vitamin D_3 and folate concentrations, other factors may also have influenced the evolution of skin color on the local scale. These possibly included: protection from xeric stress by enhanced permeability barrier function of heavily pigmented skin in arid sub-Saharan Africa (Elias et al. 2010), sexual preference for lighter-than-average skin color as originally suggested by Charles Darwin (Aoki 2002; Frost 2007; Madrigal and Kelly 2007; Madrigal and Kelly 2007), and the importance of vitamin D-mediated human antimicrobial response (Liu et al. 2006).

2.3.1. Associate examples of skin color evolution

Skin color is a polygenic trait with more than 170 currently known associated mouse genes and their human orthologs (Montoliu et al. 2012). This list includes genes involved in various cellular and physiological processes, e.g. the biosynthesis of melanin from tyrosine and cysteine, melanosome ion-trafficking, maturation and export, formation of the structural matrix and melanosome turnover (for review see Sturm and Duffy 2012). Pigmentation level is tightly controlled by a complex system of different factors such as hormones, cytokines, growth factors and their receptors, including, among others, the melanocortin 1

receptor (MC1R), its agonist α-melanocyte stimulation hormone (α-MSH) and antagonist agouti-signaling protein (ASIP) (Figure 4). UVR-induced damage increases the level of α-MSH, which upregulates the expression of brown eumelanin and enhances the photoprotective function against the damaging effects of subsequent UVR exposures by escaping UVR-induced melanocyte apoptosis (Brenner and Hearing 2008 and references therein). In contrast, the binding of the antagonist molecule reverts melanocytes to red pheomelanin production and downregulation of eumelanogenesis (Lu et al. 1994).

The *MC1R* gene was the first one to be connected with natural skin color variation among humans (Valverde et al. 1995). Mutations in this gene are generally associated with red hair, fair skin, freckling and poor tanning. The frequency of "red hair" alleles exhibits a south to north gradient in Europe and appears to be consistent with general pigmentation levels in the local populations (Gerstenblith et al. 2007). Interestingly, a study of ancient DNA suggested that approximately 1% of Neanderthals were homozygous for the loss-of-function *MC1R* allele and possibly had pale skin and/or red hair similar to that of modern humans (Lalueza-Fox et al. 2007). Furthermore, the *MC1R* gene has an uncommon variation pattern: sub-Saharan Africans possess the lowest diversity and completely lack non-synonymous substitutions, pointing to very strong functional constraints and purifying selection, while Europeans have an excess of non-synonymous variation, pointing to either relaxation of constraints (that is, the absence of strong purifying selection) (Harding et al. 2000) or balancing selection that, arguably, would have maintained enhanced genetic variability (Rana et al. 1999; Rees and Harding 2012). Traces of non-neutral evolution within the *MC1R* gene have been detected by various means in different human populations (Makova and Norton 2005): a re-sequencing-based Tajima's D approach indicates positive selection in Europe (Savage et al. 2008), while genome-wide F_{ST} data points to the selective sweep of a single non-synonymous allele in eastern Asia (Coop et al. 2009). The overall pattern of *MC1R* variation is quite unusual on the genome-wide scale, as generally autochthonous African populations have the highest nucleotide diversity level, indicating the African origin of our species (Abecasis et al. 2010).

In addition to *MC1R,* other pigmentation genes have been found to show higher than genome-wide average levels of among population differences and be enriched for signals of positive selection (Voight et al. 2006; Sabeti et al. 2007; Williamson et al. 2007; Pickrell et al. 2009). The most notable examples are: *TYR*, coding for the melanogenic enzyme tyrosinase, *SLC24A5, SLC45A2* and *OCA2*, coding for membrane transporter proteins, and *KITLG*, coding for a growth factor involved in the regulation of the number of melanocytes during development (Figure 4) (for review see: Sturm 2009; Anno et al. 2010; Rees and Harding 2012; Sturm and Duffy 2012).

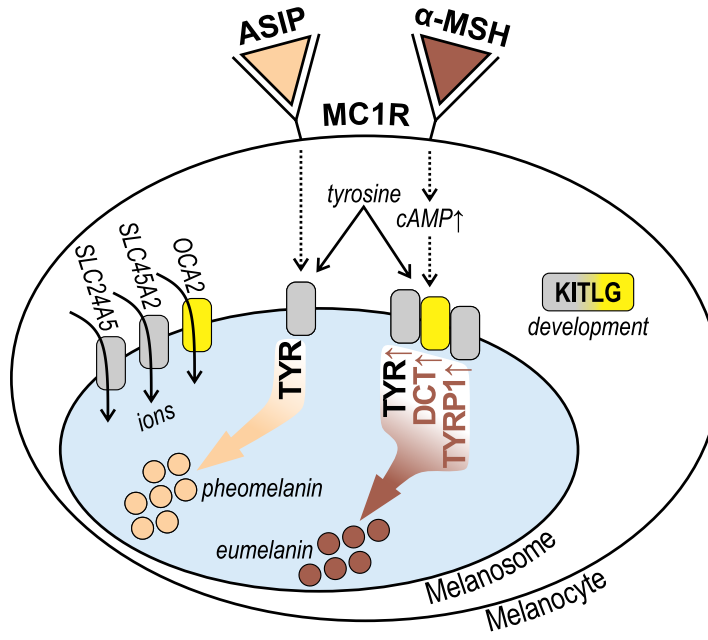

Figure 4. Schematic representation of human melanogenesis. The enzyme tyrosinase constitutes a critical rate-limiting step in melanin production from tyrosine. The stimulation of the melanocortin 1 receptor (encoded by the *MC1R* gene) by α-melanocyte stimulation hormone (α-MSH) leads to the activation of the cAMP pathway, enhanced expression of melanogenic tyrosinase (*TYR*), dopachrome tautomerase (*DCT*) and tyrosinase-related protein 1 (*TYRP1*), and the synthesis of dark eumelanin. *TYRP1* and *DCT* are only involved in eumelanin production. The antagonist of MC1R, the agouti-signaling protein (encoded by the *ASIP* gene), precludes α-MSH initiated signaling and reverts melanogenesis towards the basal state of light pheomelanin synthesis. Other important factors include melanosome transporter proteins NCKX5 (encoded by the *SLC24A5* gene), MATP (encoded by the *SLC45A2* gene) and p-protein (encoded by the *OCA2* gene), which are responsible for the transport of small molecules, ions and pH regulation. The Kit ligand (encoded by the *KITLG* gene) binds to the tyrosine receptor kinase KIT and promotes the migration, survival and proliferation of melanocytes. The *SLC24A5, SLC45A2, TYR* and *TYRP1* genes (gray shading) show traces of natural selection in the European population, while the *OCA2* and *DCT* loci (yellow shading) show traces of natural selection among East Asians. The *KITLG* gene shows traces of natural selection in both European and East Asian populations (for review see: Scherer and Kumar 2010; Sturm and Duffy 2012).

The *SLC24A5* gene variation offers probably the best example of adaptive evolution among pigmentation genes. It codes for the potassium-dependent sodium/calcium exchanger (NCKX5), essential for Ca^{2+} uptake coupled with proton transport and the regulation of melanosomal pH (Ginger et al. 2008). A single derived non-synonymous *SLC24A5* substitution (rs1426654, *p.Ala111Thr*) is

almost entirely fixed in European populations, but not in Africa and Asia (Lamason et al. 2005). Melanosomal pH level is known to be more alkaline in heavily pigmented individuals, while lightly pigmented Europeans have lower pH. Activity and/or maturation of several melanosomal proteins, including the rate-limiting tyrosinase, are optimal at neutral pH level (Fuller et al. 2001). The European-specific 111Thr allele impairs NCKX5 function, leading to the acidification of the melanosomal environment, decreased tyrosinase activity and net melanin production, which would be advantageous at higher geographical latitudes with lower UVR levels. On the contrary, optimal tyrosinase activity, which is facilitated by a fully functional NCKX5 protein and high total melanin load, is essential near the equator, as explained by the vitamin D_3/folate hypothesis (Ginger et al. 2008; Cook et al. 2009; Jablonski and Chaplin 2010). Interestingly, lightly pigmented East Asian individuals share the ancestral allele with heavily pigmented Africans (Lamason et al. 2005). This is not surprising, taking into account that anatomically modern humans dispersed independently multiple times into non-tropical latitudes of Eurasia, and, therefore, may have evolved depigmented phenotypes by means of convergent evolution (McEvoy et al. 2006; Norton et al. 2007; Jablonski and Chaplin 2010). Quantitative estimates show that the *p.Ala111Thr* amino acid change accounts for up to 38% of the European-African difference in skin melanin index (Lamason et al. 2005), and for approximately 33% of the difference between lightly and heavily pigmented individuals with South Asian ancestry (Stokowski et al. 2007). An unusually long linkage disequilibrium block and high European-specific differentiation pattern around this gene indicate a very strong recent selective sweep – an observation that has been confirmed by multiple independent studies (Norton et al. 2007; Sabeti et al. 2007; Coop et al. 2009; Pickrell et al. 2009; Grossman et al. 2013).

Another protein essential for maintaining ionic homeostasis within melanosomes is a membrane-associated transporter protein, or MATP, encoded by the *SLC45A2* gene (Dooley et al. 2013). Signs of European-specific positive selection were detected in the *SLC45A2* locus both by genotyping (Norton et al. 2007; Sabeti et al. 2007) and re-sequencing methods (Soejima et al. 2006; Grossman et al. 2013), and a strong latitudinal cline of a single non-synonymous *SLC45A2* substitution (rs16891982, *p.Leu374Phe*) was found in the European sub-continent (Lucotte et al. 2010). The derived 374Phe allele was shown to be associated with lower melanin content and lighter skin color by functional assay, supporting the importance of different transporter proteins in the regulation of melanogenesis (Cook et al. 2009).

Although two genes listed above – *SLC24A5* and *SLC45A2* – show traces of strong European adaptation, Asian-specific skin lightening alleles are known as well. For example, the non-synonymous rs1800414 mutation (*p.His615Arg*) in the *OCA2* gene is both restricted to East Asia and under selection in that region (Lao et al. 2007; Edwards et al. 2010; Yuasa et al. 2011; Donnelly et al. 2012). Other genes, including *DCT* and *TYRP1* coding for components of the melanogenic enzyme complex, as well as *ADAM17* and *ADAMTS20* similarly show

non-neutral evolution in lightly pigmented East Asians, although no causative association with pigmentation has been established yet (Norton et al. 2007; Edwards et al. 2010). Notably, *ADAM17* (UniProtKB P78536) and *ADAMTS20* (UniProtKB P59510) also play a role in physiological processes other than pigmentation (UniProt Consortium 2013 and references therein), and it is therefore possible that signatures of selection detected in these genes are due to various biological functions (Edwards et al. 2010). In addition, studies of the *TYRP1* gene coding for tyrosinase-related protein 1 indicate selection in lightly pigmented Europeans (Voight et al. 2006; Lao et al. 2007) and, possibly, Africans (Alonso et al. 2008). Interestingly, the European specific rs12913832 mutation in the *OCA2* enhancer significantly decreases the level of gene expression (Cook et al. 2009). This substitution segregates almost perfectly with blue eye color and possibly affects skin color as well (Sturm et al. 2008; Branicki et al. 2009), once again pointing to the convergent evolution of light skin in Eurasia (McEvoy et al. 2006; Norton et al. 2007). The product of the *OCA2* gene, the p-protein (UniProtKB Q04671), plays a role in the transport of melanosomal metabolites, although its exact function is not fully understood yet (UniProt Consortium 2013 and references therein).

The main rate-limiting enzyme in melanogenesis, catalyzing the first two steps and at least one subsequent step in the conversion of tyrosine to melanin, is tyrosinase, encoded by the *TYR* gene (Ray et al. 2007). Along with tyrosinase-related protein 1 (encoded by *TYRP1*) and dopachrome tautomerase (encoded by *DCT*), it forms the melanogenic enzyme complex. Several SNP genotyping studies have suggested the role of this gene in normal variation of skin pigmentation. One *TYR* non-synonymous rs1042602 (*p.Ser192Tyr*) polymorphism has been associated with differences between lightly and heavily pigmented individuals from South Asia, yet explaining only up to 2.5% difference between two cohorts (Stokowski et al. 2007). The same SNP has provided a strong signal of European-specific diversity and is associated with eye color, freckles and skin pigmentation in Europe (Shriver et al. 2003; Norton et al. 2007; Sulem et al. 2007). However, the results of scans for natural selection on this gene have been inconsistent: a few studies have detected a signal of selection (Myles et al. 2007; Norton et al. 2007; Alonso et al. 2008), while others have not been able to reject the neutral hypothesis (Izagirre et al. 2006; Voight et al. 2006; Lao et al. 2007; Sabeti et al. 2007; Williamson et al. 2007; Candille et al. 2012). The level of tyrosinase expression is similar, but enzyme activity varies among different skin color types. This can be a result of either varying melanosomal ionic environment between different phenotypes (optimal activity is present at neutral pH level), or functional differences within the enzyme itself (Fuller et al. 2001; Alaluf et al. 2003). For example, the two most common non-synonymous mutations, rs1042602 (*p.Ser192Tyr*) and rs1126809 (*p.Arg402Gln*), have a significant negative effect on enzyme activity as shown by *in vitro* analyses (Tripathi et al. 1992; Chaki et al. 2011). Therefore, additional studies are needed to investigate the patterns of diversity and selection

within the *TYR* gene and its relation with (de-)pigmentation processes among modern human populations.

The Southern route migration out of Africa implies that a group of anatomically modern humans expanded from UV-rich near-equatorial Ethiopia approximately 65 to 70 kya, moving northeastwards to South Arabia and further into Eurasia (Macaulay et al. 2005; Soares et al. 2009; Oppenheimer 2012; Soares et al. 2012; Wei et al. 2013), indicating that the ancestral proto-Eurasian population had to move and adapt to a novel UV-environment during this process. Examples of regional-specific convergent skin color evolution were discussed above, but are there any traces of evolutionary processes acting upon pigmentation genes in the proto-Eurasian population before its split into regional human populations? Miller et al. (2007) found a strong association between a non-coding substitution in the 5' flanking region of the *KITLG* gene (rs642742) and skin color, and estimated that the double-derived Eurasian allele lightens a person's color by an average of 6 to 7 melanin units in comparison to a difference of approximately 30 melanin units between heavily pigmented Africans and Europeans (mean spectrophotometrically measured melanin index is around 30 in European Americans and 58 in African Carribeans (Shriver et al. 2003)). The derived rs642742 allele is frequent among Eurasian populations, whereas the ancestral allele is virtually fixed in Africa, but not elsewhere. Previous studies have likewise highlighted this locus as a very strong candidate for selection in both Asian and European populations, therefore indicating that selection acting upon the *KITLG* gene may have been invoked in the proto-Eurasian population moving away from the UV-rich equator after the African exodus (McEvoy et al. 2006; Miller et al. 2007; Williamson et al. 2007; Coop et al. 2009; Pickrell et al. 2009).

2.3.2. A model for the evolutionary architecture of human skin pigmentation

The general pattern of variation among pigmentation genes, including melanosomal transporter proteins (*SLC24A5, SLC45A2, OCA2*) and signal transmitting receptors and ligands (*MC1R, KITLG*) agrees well with the vitamin D_3/folate hypothesis for the evolution of human skin color. It indicates that partial loss of function is tolerated or even favored by means of different evolutionary processes, that is either positive selection, balancing selection or relaxation of functional constraints, in lightly pigmented populations, to ensure lower total melanin content, decreased photoprotection and, as a result, biosynthesis of vitamin D_3 in low UV environment, such as East Asia and Europe. On the contrary, loss-of-function mutations are not tolerated in UV-rich (near-)equatorial areas, like sub-Saharan Africa, to ensure protection from UVR-induced folate photolysis (for review see: Sturm 2009; Anno et al. 2010; Rees and Harding 2012; Sturm and Duffy 2012).

The evolutionary history of human skin pigmentation is complex and may have involved several independent episodes of natural selection acting on

different genes at different time periods. Firstly, it is likely that pigmentation genes were affected by selection in the common ancestors of all living humans after they lost their protective fur (McEvoy et al. 2006) and split from other archaic hominins (Meyer et al. 2012). Secondly, selection for lighter skin acting upon the *KITLG* gene may have been invoked in the proto-Eurasian population moving away from the equator as a result of the African exodus (McEvoy et al. 2006; Lao et al. 2007; Miller et al. 2007; Williamson et al. 2007; Pickrell et al. 2009; Beleza et al. 2013). And, thirdly, selection on vitamin D_3 synthesis may have been an important evolutionary driver associated with the colonization of low UVR environments after the split of the proto-European and proto-East Asian populations 25 to 38 kya (Rasmussen et al. 2011). Several lines of evidence point to convergent evolution during the last stage of depigmentation. While the East Asians display high frequencies of population-specific alleles in pigmentation genes *OCA2, DCT, ADAM17, ADAMTS20* and in some other loci, the signals for European-specific selection are centered on a different set of genes – *SLC24A5*, *SLC45A2* and *TYRP1* (Figure 4) (Norton et al. 2007; Edwards et al. 2010; Beleza et al. 2013). Coalescence estimates suggest that the European-specific pigmentation alleles have reached high frequencies around 19 to 11 kya (Soejima et al. 2006; Beleza et al. 2013) or even later, between 12 and 3 kya (Norton and Hammer 2007). The last stage of skin depigmentation may have been driven by various factors, including climate, cultural and demographic changes during and after the Last Glacial Maximum. The decrease in winter UVR level, which peaked between 12 and 9 kya, and the extensive use of protective clothing and shelter seeking have both negatively affected vitamin D_3 bioavailability, therefore promoting further skin lightening (COHMAP 1988; Jablonski 2004). In addition, the growth of effective population size after the end of the LGM period may have been sufficient for advantageous mutations to become fixed in a relatively short time (Beleza et al. 2013), and the onset of agriculture and the associated diet shift in the early Holocene period may have contributed to the final stage of depigmentation in Europe (Cavalli-Sforza et al. 1994; Khan and Khan 2010).

The virtual lack of skin color associated genetic data from heavily pigmented populations outside of Africa, e.g. from Australasia, and the fact that the majority of available selection scans were performed on the HapMap dataset, which was until recently largely limited to only three populations from Africa (YRI), Europe (CEU) and East Asia (JPT+CHB), limits our knowledge about the evolution of human skin color on the global level. Therefore, additional follow-up re-sequencing and functional studies are needed to better understand the process of selection acting upon various pigmentation genes among different human populations and to determine how natural variation within candidate genes affects the expression and/or functional properties of proteins, i.e. to establish and explicitly explore the genotype-phenotype relationship within pigmentation associated loci.

3. AIMS OF THE STUDY

The general aim of the current study was to describe and analyze genetic variation among modern human populations relevant for the following scientific problems:

1. Settlement of prehistoric Sahul as can be seen from the variation of human haploid genomes (Refs. I and II), including:
   a. Identification of novel mtDNA and NRY lineages in endogenous populations of New Guinea and Australia and their phylogenetic and phylogeographic placement within the maternal and paternal gene pools of anatomically modern humans;
   b. Estimation of the possible time of arrival of anatomically modern humans to prehistoric Sahul;
   c. Assessment of recent gene flow to prehistoric Sahul from Eurasian populations, particularly those of South Asia, as can be deduced from human mtDNA and Y-chromosome variation.
2. Patterns of diversity and selection in the pigmentation associated human tyrosinase gene (Ref. III), including:
   a. Characterization of the worldwide phylogeny of different *TYR* haplotypes using extensive re-sequencing and genotyping of a global population sample;
   b. Assessment of signatures of natural selection acting upon the *TYR* gene using both local and genome-wide selection tests;
   c. Association of neutral genetic variation in this autosomal locus with the evolution of different skin color phenotypes in the main continental groups of anatomically modern humans.

4. MATERIALS AND METHODS

The origin of the human DNA samples analyzed in the present study, along with the experimental and analytical methods used, are described in detail in the respective research articles and/or their supporting materials.

5. RESULTS AND DISCUSSION

5.1. Settlement of prehistoric Sahul as seen from the variation of human haploid genomes (Refs. I and II)

Both genetic and archaeological data support the occupation of prehistoric Sahul by anatomically modern humans *ca* 50 kya (ref. I; Hiscock 2008). Mitochondrial DNA lineages found among Aboriginal Australian and Melanesian populations are largely unique, pointing to their autochthonous origin. This observation is further supported by the lack of recently shared NRY variation, except on the macro-haplogroup C-M130 and K-M9 level and obvious European introgressions during the very recent demographic events. The maternal gene pool of the populations of Greater Australia can be classified by the presence of mitochondrial hgs M29'Q, M42, P, S, O (referred to as N12 in ref. I) and haplotypes which are only characterized by a single complete mtDNA genome, including M14, M15 (referred to as M13 in ref. I), N13, N14, R12 and R14 (Figure 5) (ref. I and references therein). There are in total 38 complete mtDNA sequences available from Aboriginal Australian populations at present (ref. I; van Oven and Kayser 2009 and references therein), which makes them one of the least studied human groups.

All Aboriginal Australian maternal and paternal lineages fall into Eurasian mtDNA founder branches M and N, and NRY hgs C-M130 and F-M89, respectively (Figure 5). This suggests that both Australian and New Guinean populations are descendants of the same founder group of anatomically modern humans that left Africa approximately 70 kya. The coalescence dates of major mitochondrial macro-haplogroups M, N and R in the populations of prehistoric Sahul are all within the range of 53 to 58 kya (Table 1 in ref. I). These time estimates (a) support the general timeframe for Out-of-Africa expansion of anatomically modern humans, and (b) coincide with the time of arrival of anatomically modern humans to Greater Australia *ca* 50 kya as suggested by archaeological evidence. Moreover, the presence of multiple autochthonous branches of the same mitochondrial founder hg P in both Australia (P3, P4b, P5–7 and P8 (originally referred as P9 in ref. I)) and New Guinea (P1–3, P4a) most likely argues for a single founder population settling the whole region of prehistoric Sahul at least 50 kya as indicated by the deep coalescence age of the P clade (Figure 6) (ref. I). In addition, the M and N (excluding R) portions of the mtDNA phylogeny are also largely unique in both Australia (represented by hgs M42, S and O) and New Guinea (represented by hg Q), indicating their local origin and the substantial isolation of the two groups since the first colonization of Greater Australia. Notably, apart from hg P4, which likely arose close to the time of initial settlement, only two lineages are shared between northern and southern Sahul beyond macro-haplogroup M and N, and hg P level. These include hg P3 and a newly identified complete Australian mtDNA genome belonging to the Melanesian-specific hg Q (Figure 5) (ref. I). The age

of these clades was estimated to be 39,200±8,200 and 30,400±9,300 years BP (Table 1 in ref. I), respectively, indicating ancient secondary gene flow from New Guinea (Figure 6), i.e. well before the land bridge between the two regions was submerged by rising sea level approximately 8,000 years ago.

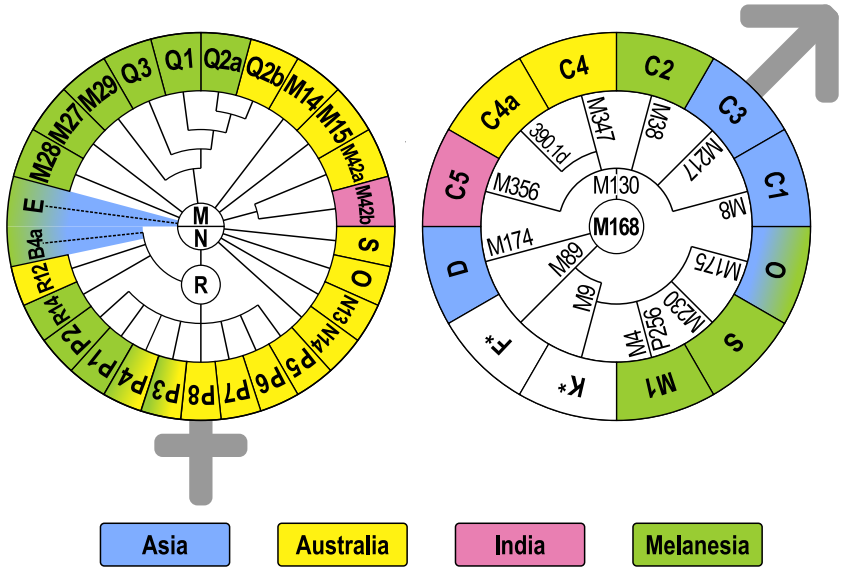

Figure 5. Simplified tree of autochthonous mtDNA and Y-chromosome haplogroups in Greater Australia (left and right spheres, respectively). All currently classified maternal branches are included. Mitochondrial haplogroup names are updated according to Build 15 (30/09/2012) of PhyloTree.org (van Oven and Kayser 2009). Haplogroup geographical origin is color-coded according to the legend. Gradient shading denotes: (a) secondary gene flow of Asian-specific mtDNA (B4a and E) and Y-chromosome (O) haplogroups into Melanesia, and (b) mtDNA sub-clades (P3 and P4) shared between populations of Australia and Melanesia. Recently identified shared common Late Pleistocene ancestry of the Aboriginal Australian M42a and South Asian M42b lineages is also shown (Kumar et al. 2009). Figure adapted from refs. I and II and updated as described above.

Similarly to mtDNA, the NRY phylogeny supports substantial pre-historic isolation and emergence of regional-specific branches (C4-M347, C4a-DYS390.1del in Australia, and C2-M38, M1-M4, S-M230 (originally referred to as K5-M230 in ref. I) in New Guinea) (Figure 5) (ref. I). Such long-standing isolation of Aboriginal Australians is in agreement with the archaeological record, which indicates only a very limited cultural contact with New Guinea islanders in the direct vicinity of Torres Strait, the historic 'Macassan' contact between populations of northwestern Australia and fishermen from the eastern

parts of Island Southeast Asia, and, of course, the more substantial recent European admixture (Mulvaney and Kamminga 1999).

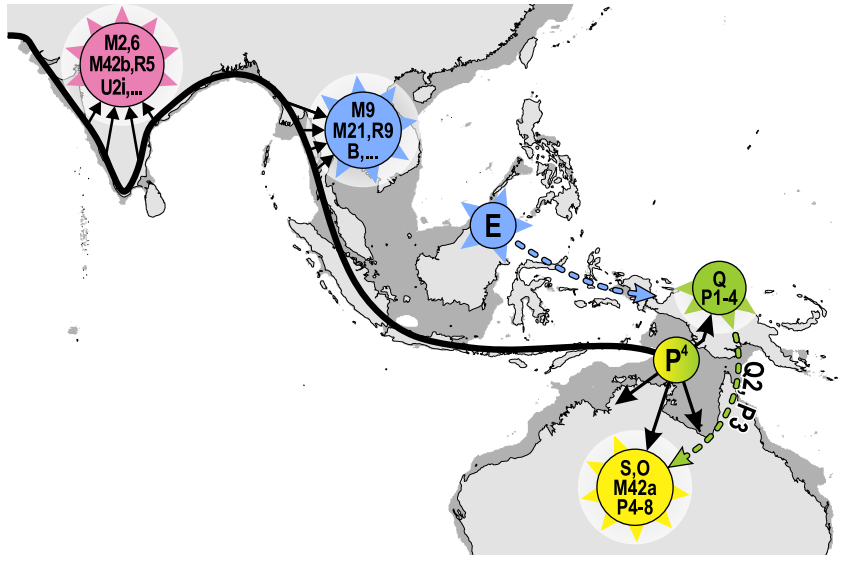

Figure 6. Peopling of Greater Australia. Map of Southeast Asia and Greater Australia, showing both the modern coastline (light gray) and the 120-m depth contour below the current sea level (dark gray), and indicating the extent of Late Pleistocene land masses. Map depicts the peopling of the continent as inferred from the extant mtDNA phylogeny. The bold black line indicates the possible southern coastal route out of Africa to Sahul. Thinner black arrows indicate possible directions of human expansions. Some of the oldest region-specific branches of the human mtDNA tree are depicted in spheres. After the initial spread of anatomically modern humans out of Africa to Sahul, the principal process is the differentiation of the mtDNA clades Q in Near Oceania and M42a, O and S in Australia. Subsequent to that process, there is little migration within Sahul other than Q and P3 from New Guinea to Australia. The genetic isolation of Australia is very clearly evident already before the land bridge between Australia and New Guinea disappears approximately 8,000 years ago. The origin of hg E on the northeastern coast of Sundaland and its mid-Holocene spread into Near Oceania are also shown (adapted from refs. I and II and updated according to Build 15 (30/09/2012) of PhyloTree.org (van Oven and Kayser 2009)).

Apart from the potential signal of secondary ancient gene flow from New Guinea, there are no other indications of genetic contacts between Aborigines and other Melanesian or Asian populations that would manifest in their mtDNA or Y-chromosome pools studied so far. More specifically, the virtual lack of mtDNA lineage sharing between South Asians and Australians, except at the

macro-haplogroup root level, contradicts the possibility of recent, mid-Holocene period gene flow from Hindustan to Sahul, as it has been argued since early studies of human demographic history (Huxley 1870). Contrary to earlier claims that Y-chromosome hg C-M130* signals recent Indian gene flow to Australia (Redd et al. 2002), here we report a new biallelic marker M347 which defines the Australian-specific hg C4 and describes all Aboriginal C-M130* Y chromosomes, including those with the common Australian-specific deletion DYS390.1. Moreover, Indian C-M130* haplotypes are now known to harbor the hg C5 defining biallelic marker M356 (Figure 5) (Sengupta et al. 2006). Therefore we suggest that Indian and Australian C-M130* Y chromosomes form a paraphyletic group that has been resolved using the new Australian-specific M347 and the published Indian-specific M356 biallelic markers. Assuming the updated resolution of NRY hg C reported here, the dissimilarity between South Asian and Australian C-M130 Y chromosomes is not consistent with the hypothesis of mid-Holocene gene flow from India to Australia (ref. I). Interestingly, a phylogenetic connection between Australian- and Indian-specific maternal hgs M42a and M42b has been recently identified and dated to approximately 55 kya, that is long before the proposed Holocene migration around 4,000 to 5,000 years BP (Kumar et al. 2009; van Holst Pellekaan 2013).

Although all Australian mtDNA and Y-chromosome samples analyzed by us belonged to haplogroups specific to Sahul, people of Near Oceania and New Guinea are known to bear significant traces of Holocene gene flow from East and insular Southeast Asia. For example, the Austronesian language migration and its genetic impact on the autochthonous Melanesian gene pool have received much broader attention than the initial Late Pleistocene settlement (for review see Kayser 2010). Apart from hg B4a sub-branches, which can be associated with the Austronesian expansion, the second most frequent non-endogenous hg in Near Oceania is hg E, a sub-clade of the Asian-specific M9 (Figure 5). This haplogroup accounts for approximately 15% of mtDNA lineages in Island Southeast Asia and Taiwan, albeit extending occasionally to the Malay Peninsula and northern Melanesia, where it is sometimes found at a considerable frequency (e.g. up to 40% in Papuan-speakers from New Britain) (Friedlaender et al. 2007; Hill et al. 2007; ref. II). We estimate the age of this haplogroup to be 33,150±8,200 years BP. Using complete mtDNA sequences, we determined that hg E has two major sub-clades, E1 and E2, coalescing around 17 and 9.5 kya, respectively. These can be further sub-divided into E1a'b and E2a'b (Figure 2 in ref. II). The geographical distribution, diversity and age estimates of different E sub-clades suggest their *in situ* origin in Island Southeast Asia, either on the northeastern coast of Sundaland or in the neighboring northwestern Wallacea (northern Sulawesi) (Figure 3 in ref. II), whereas the unusually high frequency of hg E and its spotty distribution in Near Oceania can be explained by the extreme level of genetic drift among insular populations (Friedlaender et al. 2007). Coalescence age estimates of the two most frequent sub-clades, E1a and E2, suggest that the expansion of hg E started in early Holocene, around 11,000–12,000 years BP, and introduced its sub-types north-

wards into Taiwan roughly 4,000–8,000 years ago, and eastwards into Near Oceania (ref. II). The latter region was probably reached about 5,000 year BP or slightly later (Friedlaender et al. 2007). It is possible that hg E expansion was triggered by rising sea level after the end of the Last Glacial Maximum: the inundation of the Sunda shelf forced the displacement of coastal-dwelling populations into new habitats and, therefore, the spread of hg E across Island Southeast Asia and northern Melanesia (Hill et al. 2007; ref. II).

5.2. Global patterns of diversity in the pigmentation associated tyrosinase gene (Ref. III)

Tyrosinase is the main rate-limiting melanogenic enzyme that plays a pivotal role in the conversion of tyrosine to melanin. We have studied worldwide patterns of variability in the pigmentation associated *TYR* gene, using three different sample sets and analytical methods: (a) re-sequencing of approximately 24.3 kb, including complete coding, partial intronic, partial 5' and 3' flanking regions, was performed in 81 human DNA samples from six major continental groups (Africa, America, Europe, South Asia, East Asia and Oceania) and complemented by an array of sequence-based selection tests, (b) Illumina genotyping data from 351 samples from the same six populations were used to study patterns of selection using whole-genome haplotype length (iHS and XP-EHH) and population differentiation F_{ST}-based approaches, and (c) Illumina *TYR* genotyping data from 1108 individuals, including the HGDP-CEPH Human Genome Diversity Project panel, were analyzed using an ancestral recombination graph in order to reconstruct and visualize phylogenetic relationships within our sample. To increase phylogenetic resolution, the latter analysis was complemented by re-sequencing of additional informative SNPs and was performed using seven population groups: Africa was sub-divided into northern (represented by the HGDP-CEPH Mozabite sample) and sub-Saharan regions. Phasing of individual singleton mutations during the re-sequencing project was performed *in vitro* by cloning. This allowed us to improve the assignment of singletons to individual chromosomes and, therefore, to obtain unbiased coalescence ages estimates.

The highest coding nucleotide diversity in our re-sequencing panel was detected in the European sample (0.00067), followed by Africans (0.00021). The non-synonymous variation within the *TYR* gene was largely restricted to only two SNPs, rs1042602 (*p.Ser192Tyr*) in exon 1 and rs1126809 (*p.Arg402Gln*) in exon 4. Both mutations are known to significantly affect enzymatic activity of tyrosinase *in vitro* and are largely conserved among different species (Tripathi et al. 1992; Chaki et al. 2011; ref. III). Three additional non-synonymous polymorphisms that were detected have low frequency: two are singletons (one in the African and another in the European sample) and one is a doubleton found in Europe. The ancestral recombination graph shows that the majority of non-synonymous polymorphisms are centered in Europe or

North Africa: 72% of total non-synonymous variation was confined to these two populations (Figure 3 in ref. III). The high incidence of European-specific non-synonymous alleles in northern Africa is not surprising. North Africans derive from multiple founders, including those with recent European descent: the amount of European-specific ancestry in HGDP-CEPH Mozabite sample was estimated around 80% (Price et al. 2009). Interestingly, Mozabites are also characterized by lightly pigmented skin. Therefore, allele sharing between Europe and North Africa reflects common demographic history, rather than convergent evolution.

We also found a region 6.9 kb upstream from the first codon position that shows the lowest diversity in all six continental groups studied (Figure 1 in ref. III). As no regulatory motifs have so far been described in this region to our knowledge, our analyses highlight it as a new potential candidate of functional importance.

The observed pattern of variation within the *TYR* gene can be summarized as follows. Firstly, Europeans have the highest coding nucleotide diversity (Figure 1 in ref. III). In contrast, on a genome-wide scale, populations from the African continent usually possess higher diversity estimates. Secondly, sub-Saharan Africans almost totally lack non-synonymous alleles in the tyrosinase gene likely because of purifying selection (Figure 3 in ref. III). Thirdly, lightly pigmented East Asian individuals are virtually depleted for non-synonymous variants (Figure 3 in ref. III). Genetic diversity within both lightly pigmented Europeans and heavily pigmented sub-Saharan Africans can be explained by the general evolutionary model of human skin pigmentation: mutations leading to decreased photoprotection will be deleterious near the equator due to extensive folate photolysis. In contrast, polymorphisms of this kind will be well tolerated due to relaxed pressure of purifying selection at higher UV-poor latitudes. Furthermore, mutations depressing tyrosinase functional activity and, therefore, allowing for enhanced vitamin D_3 synthesis might be even favored in this environment. Interestingly, a similar pattern of diversity has previously been observed in pigmentation associated *MC1R* loci (Rana et al. 1999; Harding et al. 2000). The lack of non-synonymous allele sharing between lightly pigmented Europeans and East Asians should not be surprising (Figure 3 in ref. III). There is a substantial amount of evidence supporting convergent evolution of skin depigmentation in Eurasia: *SLC24A5, SLC45A2, TYRP1* and, as we also show here, *TYR* alleles are associated with skin lightening in Europe, while *OCA2, DCT* and some other genes are involved in skin depigmentation in East Asia (Figure 4) (Norton et al. 2007; Edwards et al. 2010; Beleza et al. 2013).

Detected positive Tajima's D test value indicates possible traces of balancing selection in the *TYR* 5' flanking region and exon 1 in our European sample (Table 1 in ref. III). However, European D statistics is only significant before the correction for multiple testing, and thus should be treated as suggestive. The significance of Tajima's D selection tests was assessed by comparison to coalescent simulations under the *best-fit* model in the COSI software, which has been calibrated to reproduce extant human genetic variation under

neutrality (Schaffner et al. 2005). Nevertheless, other processes, e.g. population sub-divisions, can likewise yield positive D values and, although no differentiation was detected within our European *TYR* sample, caution is needed while interpreting the observed result. Several lines of evidence, including highest incidence of non-synonymous alleles in populations with European descent, positive Tajima's D estimates and at least threefold higher coding nucleotide diversity in Europe than in any other studied sample, imply that variation in the *TYR* gene has potentially played an important role in the adaptation to low UV environment in European subcontinent. Relaxed selective constraint in this UV-poor geographical region can be put forward as the most parsimonious explanation for the detected pattern of tyrosinase genetic variation, a process which has been previously described in the *MC1R* locus (Harding et al. 2000). The role of balancing selection in the evolution of the *TYR* gene must be further examined using larger sample size and/or complete gene re-sequencing strategy.

Significantly positive Tajima's D values were similarly observed in studied Oceanian group in both 5' flanking region and exon 1, and complete re-sequencing alignment. However, taking into the account complex demographic history of Oceanian population, which included small founding population, severe bottlenecks and long-term isolation as described in the first part of the present thesis, interpretation of these results should be considered with caution. Furthermore, additional bias may have been introduced by the lack of the specific parameters of Oceanian demographic history in the *best-fit* model, genetic heterogeneity and relatively small size of studied Oceanian sample (n=14 chromosomes), which included Papuans and Bougainville Islanders.

In addition to sequence-based demography-corrected Tajima's D test, haplotype length-based tests looking for longer than average span of linkage disequilibrium were employed. None of them reflected any deviation from neutrality (Table 2 in ref. III). This approach operates at a shallow evolutionary time depth, since recombination will quickly degrade long haplotype blocks produced during selective sweeps. Therefore, inconsistency between different tests for selection is not surprising and has already been reported earlier (de Gruijter et al. 2011). Still, our European sample had the highest population differentiation F_{ST}-based and cross-population extended haplotype homozygosity (XP-EHH) scores, belonging to the top 71th and 87st percentiles of genome-wide distribution, respectively.

Human *TYR* phylogenetic tree coalesces between 1.4 to 1.6 MYA, i.e. close to the mean coalescence time of human autosomal genes (Blum and Jakobsson 2011). We further calculated the age of the two monophyletic clusters defined by non-synonymous substitutions rs1042602 (*p.Ser192Tyr*, defining hg B1) and rs1126809 (*p.Arg402Gln*, defining hg C2a), using rho statistics and median-joining network of *TYR* re-sequencing alignment (Figure S4 in ref. III), and Bayesian coalescent approach implemented in BEAST (Drummond et al. 2012). Both alleles are common in Europe and North Africa. The first and the most frequent cluster, referred to as hg B1 in the ancestral recombination graph

(Figure 3 in ref. III), coalesces at 6,100±3,600 years BP as estimated by rho and 15,600 years BP (95%CI=400–46,600) as estimated by Bayesian coalescent approach. The second one, hg C2a, coalesces at 20,400±13,600 and 29,400 (95%CI=1,900–68,600) years BP as estimated by rho and Bayesian coalescent approach, respectively. The discrepancies between rho- and Bayesian-based estimates could be largely contributed to the different parameter settings of the two computation methods. The coalescence age of both non-synonymous substitutions postdate the split of Western and Eastern Eurasian populations between 25,000 to 38,000 years BP (Rasmussen et al. 2011) supporting the convergent evolution of light skin in Europe and East Asia. Moreover, although the mean estimates of these statistics should be treated with caution, the most recent common ancestors of hgs B1 and C2a appear to, at least partly, overlap with the age of expansion of other European-specific skin depigmentation alleles in the *SLC24A5, SLC45A2* and *TYRP1* loci, which has been dated by various authors as between 19,000 and 11,000 years BP, or between 12,000 and 3,000 years BP (Soejima et al. 2006; Norton and Hammer 2007; Beleza et al. 2013). Therefore, it is possible that the genetic variation within the *TYR* gene and its 5' flanking region has contributed to the same adaptive process in the ancestry of the European population as the *SLC24A5, SLC45A2* and *TYRP1* pigmentation loci. This stage of European-specific adaptation may have been driven by different climatic and socio-cultural processes during and after the Last Glacial Maximum and was likely complemented by convergent skin depigmentation in East Asia.

6. CONCLUSIONS

Settlement of prehistoric Sahul as seen from the variation of human haploid genomes (Refs. I and II):

1. All matrilineal and patrilineal genetic variation in prehistoric Sahul derives from the same founders (macro-haplogroups M and N in mtDNA, and C-M130 and F-M89 in the Y-chromosome tree) as other non-African populations.
2. Inferred coalescence age estimates of human mtDNA variation in Greater Australia concur with the time frame attributed to the Out-of-Africa expansion and archaeological evidence for the first human occupation of prehistoric Sahul.
3. The distribution and age estimates of different mtDNA haplogroups are consistent with the model implying a single early settlement of prehistoric Sahul approximately 50,000 years ago.
4. The lack of common mtDNA branches with other Eurasian populations (except those on the macro-haplogroup level) points to a considerable isolation of Australia and New Guinea after the initial settlement.
5. Limited mtDNA haplogroup sharing and the absence of shared NRY lineages among Australian and New Guinean populations imply rapid segregation of a single founder population to proto-Australians and proto-Melanesians.
6. There are only two potential signals of ancient secondary migration from New Guinea to Australia, reflected by the spread of matrilineal types P3 and Q2.
7. Analysis of Aboriginal Australian mtDNA and NRY variation did not reveal any traces of prehistoric mid-Holocene gene flow from South Asia to Australia.
8. Populations from New Guinea and northern Melanesia have traces of mid-Holocene gene flow from Island Southeast Asia, which is reflected by the spread of mitochondrial hg E.
9. Phylogenetic resolution of maternal hg E has been refined here using complete mtDNA sequences. This haplogroup likely arose approximately 33 kya in Island Southeast Asia and started to expand around 12 kya. The expansion may have been driven by rising sea level after the end of the Last Glacial Maximum.

Global patterns of diversity in the pigmentation associated tyrosinase gene (Ref. III):

1. The global phylogenetic tree of the human *TYR* locus was reconstructed using re-sequencing and genotyping approaches. Different sequence-based and whole-genome genotyping-based selection tests were employed to assess the patterns of natural selection acting upon the *TYR* gene.

2. The highest coding nucleotide diversity was found in the European sample, which also had the highest number of non-synonymous variants among all continental groups studied. Sub-Saharan Africans and East Asians virtually lack derived amino-acid variants.
3. Observed geographic patterns of *TYR* variation support the general evolutionary model of human skin pigmentation. This model predicts that non-synonymous mutations in pigmentation genes leading to diminished protein function and photoprotection will be deleterious near the equator. On the contrary, such variants will be tolerated due to the relaxed functional constraints in low UV environments, e.g. Europe. The virtual lack of *TYR* non-synonymous alleles in East Asia is concordant with previous studies and provides additional evidence in support of convergent evolution of light skin in Eurasia.
4. Excess of non-synonymous alleles, which profoundly depress tyrosinase functional activity, along with the highest coding nucleotide diversity detected in our European sample, suggest that relaxation of functional constraints has shaped the pattern of *TYR* variation in the Europe.
5. Two common non-synonymous variants, rs1042602 and rs1126809, in our worldwide sample are largely restricted to populations with European descent. The coalescence age of both alleles postdate the split between Western and Eastern Eurasian populations and is consistent with the age of European-specific sweeps in the *SLC24A5, SLC45A2* and *TYRP1* pigmentation associated loci.

7. SUMMARY IN ESTONIAN

Mitokondriaalse DNA, Y-kromosoomi ja türosinaasi varieeruvus Euraasia ja Okeaania populatsioonides

Inimese evolutsiooni erinevate tahkude uurimisel on tema tänase geneetilise mitmekesisuse väljaselgitamine olnud väga oluline. Valdav enamus meie olemasolevatest teadmistest kaasaegse inimese rännuteede ja tema geneetilise varieeruvuse kohta tänaste inimpopulatsioonide sees ja vahel on saadud tänu haploidsetele geneetiliste markerite uurimisele – mitokondri DNA (mtDNA) ja Y-kromosoomi varieeruvusandmetele tuginedes. Alles päris hiljuti on seda andmestikku hakanud täiendama ka kogu inimgenoomi järjendamise alusel tehtud populatsioonigeneetilised tööd. Mitokondri ja Y-kromosoomi DNA-l on mõned ainulaadsed omadused, mis annavad neile teiste geneetiliste markeritega võrreldes evolutsiooniprotsessi uurimisel olulise eelise: a) põlvnemine ainult üht vanemliini pidi, b) rekombineerumise puudumine (v.a. väike osa Y-kromosoomist) ja c) mitokondri genoomi teatud piirkondade suhteliselt suur molekulaarse evolutsiooni kiirus muu inimgenoomiga võrreldes. Nende kahe haploidse markeri varieeruvuse uurimine maailma erinevais paigus täna elavates populatsioonides on meil aidanud mõista anatoomiliselt kaasaegse inimese demograafilist ajalugu – tema väljarännet Aafrika algkodust ja muu maailma asustamist peale seda.

Käesoleva töö esimeses osas käsitletakse Austraalia ja Uus-Guinea põliselanike eelajalooga seotut läbi haploidsete andmete analüüsi. Selle piirkonna vanimad inimfossiilid on dateeritud 50,000 aasta vanuseks ja pärinevad ajast, mil Austraalia ja Uus-Guinea koos Tasmaania ja paljude väikesaartega moodustasid ühtse mandri – Sahuli. Austraalia aborigeenidele ja uus-guinealastele ühise mtDNA haplogrupi P levikumuster viitab sellele, et nii Austraalia kui Uus-Guinea asustati kas ühe migratsioonilainega või siis mitmest geneetiliselt sarnasest lähtepopulatsioonist. Emaliinide fülogeneesipuu paikse tekkega sügavate harude esinemine mõlemas eelajaloolise Sahuli populatsioonis viitab nende täielikule isoleeritusele peale algset asustamist. Lisaks sellele peetakse tõestatuks, et Uus-Guinea rahvastikku on mõjutanud ka kaks hiljutist Aasiast lähtunud migratsioonilainet. Nendest viimane, umbes 5,500 aastat tagasi ilmselt Taiwanilt lähtunud uustulnukate sissevool tõi kaasa austroneesia algupäraga geenid, samas varasem, umbes 12,000 aastat tagasi Kagu-Aasia saarestikust alanud sisseränne aga nn. pre-austroneesia geenid. Võimalik hiljutise geneetilise ühisosa temaatika Austraalia aborigeenide ja teiste aasialaste vahel on olnud tuliste vaidluste pärusmaa. Juba 1870 aastal tõi Thomas Huxley esile Austraalia aborigeenide ja mõnede India rahvaste märgatava sarnasuse. Hiljem pakuti selle põhjusena välja holotseeni keskel toimunud rahvastiku rännet Lõuna-Aasiast Austraaliasse. Ehkki see hüpotees ei ole kaasaegsete Austraalia arheoloogide seas suurt poolehoidu leidnud, on selle tõestuseks välja toodud India ja Austraalia Y-kromosoomi haplogrupi C-M130* sarnaste harude olemasolu.

Käesoleva töö esimese osa põhiline eesmärk oli uuendada meie teadmisi kahe eelajaloolise Sahuli populatsiooni geneetilise varieeruvuse kohta, kasutades selleks mtDNA täisjärjestuste ja Y-kromosoomi järjestuste fülogeneetilist analüüsi. Selleks sekveneeriti täiendavalt uute indiviidide DNA nimetatud markerite osas ja tulemusi analüüsiti koos kirjanduses avaldatud andmestikuga. Saadud andmed inkorporeeriti inimese ülemaailmse valimi põhjal koostatud ema- ja isaliinides päranduva varieeruvuse alusel rekonstrueeritud mtDNA ja Y-kromosoomi fülogeneesipuudesse. Töö tulemustest lähtuvalt võib välja tuua järgmised olulised punktid:

1. Austraalia ja Uus-Guinea tänaste põliselanike mtDNA varieeruvus on välja kasvanud samadest asutaja-makrohaplogruppidest M ja N, millest peaaegu kõik teised Euraasias levinud emaliinidki. Erinevate mtDNA haplogruppide levikumuster ja arvatavad ekspansiooniajad on kooskõlas mudeliga, mille kohaselt Sahuli asustamine leidis aset ühekordse varajase asustamislainena umbes 50,000 aasta eest.
2. Mõlemas uuritud regioonis on levinud mitmed paikse tekkega sügavad emaliinide fülogeneesipuu harud: M42a, S, O, P4b ja P6 Austraalias, P1, P2, P4a ja Q Uus-Guineas. Nende liinide puudumine Euraasias viitab mõlema uuritud regiooni valdavale isoleeritusele pärast algset asustamislainet.
3. Uue, Austraaliast leitud Melaneesia-spetsiifilise haplogrupi Q2 haru vanus näitab, et geenisiire kahe eelajaloolise Sahuli populatsiooni vahel on toimunud väga ammu ja leidis aset juba enne Austraalia ja Uus-Guinea vahelise maasilla kadumist umbes 8,000 aasta eest. Seda hüpoteesi toetab ka sügav ajaline lõhe Austraalias ja Uus-Guineas levinud haplogrupi P3 fülogeneesipuu harude vahel.
4. Uus-Guinea populatsioonide geenitiiki on mõjutanud nii Taiwanilt lähtunud austroneesia kui ka Kagu-Aasia saartelt lähtunud geenide sissevool, viimane neist tõi endaga kaasa mtDNA haplogrupp E liinid. See haplogrupp on ilmselt tekkinud eelajaloolise Sundamaa – ühise Kagu-Aasiat ning Sumatra, Borneo, Jaava ja Bali saari ning paljusid väikesaari ühendanud maamassi – põhjarannikul umbes 33,000 aastat tagasi. Haplogrupp E liinide põhja- ja idasuunalise leviku Melaneesia põhjaosa ja Taiwani suunas vallandas ilmselt viimase jääaja järgne merepinna tõus.
5. Meie leidsime uue Y-kromosoomi bialleelse markeri M347, mis määratleb Austraalia-spetsiifilise Y-haplogrupi C4 ja mille alusel saab kirjeldada kõiki Austraalia põliselanike seas levinud C-M130* tüüpi Y-kromosoome. Veelgi enam, praegu teatakse, et India C-M130* haplotüübid kuuluvad haplogruppi C5, mida defineerib bialleelne marker M356. Sellele toetudes pakume me välja, et India ja Austraalia C-M130* Y-kromosoomid moodustavad parafüleetilise grupi ning seetõttu võib väita, et Lõuna-Aasia ja Austraalia C-M130 Y-kromosoomid on erinevad ja nende levik ei ole kooskõlas oletatava, holotseenis aset leidnud geenivoo hüpoteesiga Indiast Austraaliasse.

Teine osa tööst käsitleb pigmentatsiooniga seotud türosinaasi (*TYR*) lookuse geneetilist varieeruvust. Inimese nahavärv on polügeenne tunnus ja on põhiliselt

määratud erinevat tüüpi melaniini – tumeda, valguskaitset pakkuva eumelaniini ja heleda feomelaniini – osakaalu ja jaotusega nahas. Inimese nahavärvuse ülemaailmset jaotust arvestades on üldlevinud seisukohaks, et meie nahk peab ekvaatori lähedal olema piisavalt tume selleks, et kaitsta meid kahjuliku UV-kiirguse ja selle poolt indutseeritava folaatide lagunemise eest, kuid piisavalt hele suurematel geograafilistel laiuskraadidel, näiteks Euroopas ja Ida-Aasias, et võimaldada efektiivset vitamiin D_3 sünteesi. Globaalset nahavärvi varieeruvust mõjutanud evolutsiooniprotsessid on olnud komplekssed, pigmentatsioonigeenide adaptiivse evolutsiooni kohta võib tuua mitmeid näiteid. On näiteks selgeid tõendeid, et Euraasia eri paigus esinev hele nahavärv on tekkinud konvergentse evolutsiooni tulemusena. Kui Ida-Aasias on kõrge sagedusega levinud kindlad pigmentatsioonigeenide *OCA2, DCT, ADAM17, ADAMTS20* ja veel mõne muu lookuse populatsioonispetsiifilised variandid, siis Euroopas on loodusliku valiku adaptiivsele toimele olnud allutatud hoopis teised geenid – *SLC24A5, SLC45A2, TYRP1*, ja nagu me selles töös näitame, *TYR*. Liiati on tõendeid heledamale nahavärvile viinud positiivse valiku mõjust *KITLG* geenile ida-aasialaste ja eurooplaste eellaspopulatsioonis. Töö teise osa põhiliseks eesmärgiks on kirjeldada türosinaasi, üht põhilist melanogeenset ensüümi kodeeriva *TYR* lookuse ja sellega piirneva 5' regulaatorala loomulikku mittepatogeenset geneetilist varieeruvust. Kokkuvõtvalt võib välja tuua järgmised tulemused:

1. Resekveneerimise ja genotüpeerimise tulemuste alusel koostati inimese *TYR* lookuse globaalne fülogeneesipuu. *TYR* geeni mõjutanud loodusliku valiku rolli määramiseks kasutati erinevaid geenijärjestuste ja kogu genoomi varieeruvusandmestikul põhinevaid valikuteste.
2. Suurim *TYR* lookuse kodeeriva ala nukleotiidne diversiteet leiti Euroopa valimist, mida iseloomustas ka suurim mittesünonüümsete asendustega variantide arv kõikide uuritud kontinentaalsete gruppide seas. Sahara-tagune Aafrika ja Ida-Aasia elanikel mittesünonüümsete asendustega variandid peaaegu puudusid.
3. *TYR* alleelide geograafiline levikumuster toetab inimese naha pigmentatsiooni üldist evolutsioonilist mudelit. See mudel eeldab, et pigmentatsioonigeenides toimunud mittesünonüümsed asendused viivad ensüümi aktiivsuse ja valguskaitse vähenemisele, selliste asendustega variandid eemaldatakse ekvaatori lähedal asuvatest populatsioonidest negatiivse looduliku valiku poolt. Vastupidiselt sellele on mittesünonüümsed asendused sallitud madala UV-kiirgusega keskkonnas – nagu seda on näiteks Euroopa – või loodusliku valiku poolt isegi soositud, sest võimaldavad efektiivset vitamiin D_3 sünteesi. Mittesünonüümsete asendustega *TYR* alleelide puudumine või väga väike esinemissagedus Ida-Aasias on kooskõlas varasemate uuringutega ja kallutab vaekaussi veelgi Euraasia heleda nahavärvi konvergentse evolutsiooni kasuks.
4. Euroopa valimis täheldatud türosinaasi aktiivsust märgatavalt pärssivate mittesünonüümsete alleelide suur hulk koos kodeeriva ala suurima nukleotiidse mitmekesisusega viitab sellele, et Euroopas on *TYR* varieeruvuse

mustrit võinud mõjutada funktsionaalsete piirangute nõrgenemine sellele geenile suurematel laiuskraadidel.

5. Meie ülemaailmse valimi kaks mittesünonüümsete asendustega *TYR* lookuse varianti, rs1126809 ja rs1042602, on levinud peamiselt Euroopa päritolu populatsioonides. Keskmine türosinaasi geeni haplogruppide koalestseerumise aeg, mis nende asenduste kaudu määrati, on vastavalt 29,400 või 20,400 ning 15,600 või 6,100 aastat, sõltuvalt kasutatud arvutamise meetodist. Need ajamäärangud jäävad perioodi, mil Lääne- ja Ida-Euraasia populatsioonid olid juba lahknenud ning langevad ligilähedaselt kokku ajaga, mil arvatavalt toimis Euroopa-spetsiifiline valikuline "luuatõmme" pigmentatsiooniga seotud lookustele *SLC24A5, SLC45A2* ja *TYRP1* – see on erinevate autorite poolt dateeritud kas 19,000 kuni 11,000 aasta või siis 12,000 kuni 3,000 aasta tagusesse perioodi. Selle Euroopale omase kohastumuse teke võis olla tingitud erinevatest klimaatilistest ja sotsiaal-kultuurilistest protsessidest viimase jääaja maksimumi lõppjärgus ja seda täiendas konvergentselt Ida-Aasia populatsioonide naha depigmentatsioon.

8. REFERENCES

Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME and McVean GA. 2010. A map of human genome variation from population-scale sequencing. *Nature*. 467: 1061–1073.

Alaluf S, Barrett K, Blount M and Carter N. 2003. Ethnic variation in tyrosinase and TYRP1 expression in photoexposed and photoprotected human skin. *Pigment Cell Res*. 16: 35–42.

Albrechtsen A, Nielsen FC and Nielsen R. 2010. Ascertainment biases in SNP chips affect measures of population divergence. *Mol Biol Evol*. 27: 2534–2547.

Alonso S, Izagirre N, Smith-Zubiaga I, Gardeazabal J, Diaz-Ramon JL, Diaz-Perez JL, Zelenika D, Boyano MD, Smit N and de la Rua C. 2008. Complex signatures of selection for the melanogenic loci TYR, TYRP1 and DCT in humans. *BMC Evol Biol*. 8: 74.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature*. 290: 457–465.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM and Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. 23: 147.

Anno S, Ohshima K and Abe T. 2010. Approaches to understanding adaptations of skin color variation by detecting gene-environment interactions. *Expert Rev Mol Diagn*. 10: 987–991.

Aoki K. 2002. Sexual selection as a cause of human skin colour variation: Darwin's hypothesis revisited. *Ann Hum Biol*. 29: 589–608.

Ardalan A, Oskarsson M, Natanaelsson C, Wilton A, Ahmadian A and Savolainen P. 2012. Narrow genetic basis for the Australian dingo confirmed through analysis of paternal ancestry. *Genetica*. 140: 65–73.

Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli EL, Silva NM, Kivisild T, Torroni A and Villems R. 2012. A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *Am J Hum Genet*. 90: 675–684.

Behar DM, Villems R, Soodyall H, Blue-Smith J, Pereira L, Metspalu E, Scozzari R, Makkan H, Tzur S, Comas D, et al. 2008. The dawn of human matrilineal diversity. *Am J Hum Genet*. 82: 1130–1140.

Beleza S, Santos AM, McEvoy B, Alves I, Martinho C, Cameron E, Shriver MD, Parra EJ and Rocha J. 2013. The timing of pigmentation lightening in Europeans. *Mol Biol Evol*. 30: 24–35.

Betty DJ, Chin-Atkins AN, Croft L, Sraml M and Easteal S. 1996. Multiple independent origins of the COII/tRNA(Lys) intergenic 9-bp mtDNA deletion in aboriginal Australians. *Am J Hum Genet*. 58: 428–433.

Blum MG and Jakobsson M. 2011. Deep divergences of human gene trees and models of human origins. *Mol Biol Evol*. 28: 889–898.

Bowler JM, Johnston H, Olley JM, Prescott JR, Roberts RG, Shawcross W and Spooner NA. 2003. New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature*. 421: 837–840.

Branda RF and Eaton JW. 1978. Skin color and nutrient photolysis: an evolutionary hypothesis. *Science*. 201: 625–626.

Branicki W, Brudnik U and Wojas-Pelc A. 2009. Interactions between HERC2, OCA2 and MC1R may influence human pigmentation phenotype. *Ann Hum Genet*. 73: 160–170.

Brenner M and Hearing VJ. 2008. The protective role of melanin against UV damage in human skin. *Photochem Photobiol*. 84: 539–549.

Brown P. 2013. Palaeoanthropology: Of humans, dogs and tiny tools. *Nature*. 494: 316–317.

Brown WM. 1980. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc Natl Acad Sci U S A*. 77: 3605–3609.

Candille SI, Absher DM, Beleza S, Bauchet M, McEvoy B, Garrison NA, Li JZ, Myers RM, Barsh GS, Tang H, et al. 2012. Genome-wide association studies of quantitatively measured skin, hair, and eye pigmentation in four European populations. *PLoS ONE*. 7: e48294.

Cann RL, Stoneking M and Wilson AC. 1987. Mitochondrial DNA and human evolution. *Nature*. 325: 31–36.

Cavalli-Sforza LL, Menozzi P and Piazza A. 1994. *The history and geography of human genes*. Princeton, N.J., Princeton University Press.

Chaki M, Sengupta M, Mondal M, Bhattacharya A, Mallick S, Bhadra R and Ray K. 2011. Molecular and functional studies of tyrosinase variants among Indian oculocutaneous albinism type 1 patients. *J Invest Dermatol*. 131: 260–262.

Chandrasekar A, Kumar S, Sreenath J, Sarkar BN, Urade BP, Mallick S, Bandopadhyay SS, Barua P, Barik SS, Basu D, et al. 2009. Updating phylogeny of mitochondrial DNA macrohaplogroup m in India: dispersal of modern human in South Asian corridor. *PLoS ONE*. 4: e7447.

Chaplin G. 2004. Geographic distribution of environmental factors influencing human skin coloration. *Am J Phys Anthropol*. 125: 292–302.

COHMAP. 1988. Climatic changes of the last 18,000 years: observations and model simulations. *Science*. 241: 1043–1052.

Cook AL, Chen W, Thurber AE, Smit DJ, Smith AG, Bladen TG, Brown DL, Duffy DL, Pastorino L, Bianchi-Scarra G, et al. 2009. Analysis of cultured human melanocytes based on polymorphisms within the SLC45A2/MATP, SLC24A5/NCKX5, and OCA2/P loci. *J Invest Dermatol*. 129: 392–405.

Coop G, Pickrell JK, Novembre J, Kudaravalli S, Li J, Absher D, Myers RM, Cavalli-Sforza LL, Feldman MW and Pritchard JK. 2009. The role of geography in human adaptation. *PLoS Genet*. 5: e1000500.

Corser CA, McLenachan PA, Pierson MJ, Harrison GL and Penny D. 2012. The Q2 mitochondrial haplogroup in Oceania. *PLoS ONE*. 7: e52022.

Cruciani F, Trombetta B, Massaia A, Destro-Bisol G, Sellitto D and Scozzari R. 2011. A revised root for the human Y chromosomal phylogenetic tree: the origin of patrilineal diversity in Africa. *Am J Hum Genet*. 88: 814–818.

Darwin C. 1871. *The descent of man, and selection in relation to sex*. New York, D. Appleton and company.

de Gruijter JM, Lao O, Vermeulen M, Xue Y, Woodwark C, Gillson CJ, Coffey AJ, Ayub Q, Mehdi SQ, Kayser M, et al. 2011. Contrasting signals of positive selection in genes involved in human skin-color variation from tests based on SNP scans and resequencing. *Investig Genet*. 2: 24.

Donnelly MP, Paschou P, Grigorenko E, Gurwitz D, Barta C, Lu RB, Zhukova OV, Kim JJ, Siniscalco M, New M, et al. 2012. A global view of the OCA2-HERC2 region and pigmentation. *Hum Genet*. 131: 683–696.

Dooley CM, Schwarz H, Mueller KP, Mongera A, Konantz M, Neuhauss SC, Nusslein-Volhard C and Geisler R. 2013. Slc45a2 and V-ATPase are regulators of melanosomal pH homeostasis in zebrafish, providing a mechanism for human pigment evolution and disease. *Pigment Cell Melanoma Res*. 26: 205–217.

Drummond AJ, Suchard MA, Xie D and Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 29: 1969–1973.

Durazo-Arvizu RA, Aloia JF, Dugas LR, Tayo BO, Shoham DA, Bertino AM, Yeh JK, Cooper RS and Luke A. 2013. 25-hydroxyvitamin D levels in African American and Nigerian women. *Am J Hum Biol*. 25: 560–562.

Edwards M, Bigham A, Tan J, Li S, Gozdzik A, Ross K, Jin L and Parra EJ. 2010. Association of the OCA2 polymorphism His615Arg with melanin content in east Asian populations: further evidence of convergent evolution of skin pigmentation. *PLoS Genet*. 6: e1000867.

Elias PM, Menon G, Wetzel BK and Williams J. 2010. Barrier requirements as the evolutionary "driver" of epidermal pigmentation in humans. *Am J Hum Biol*. 22: 526–537.

Endicott P and Ho SY. 2008. A Bayesian evaluation of human mitochondrial substitution rates. *Am J Hum Genet*. 82: 895–902.

Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H and Brinkmann B. 1998. Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. *Mol Biol Evol*. 15: 1108–1114.

Forster P, Torroni A, Renfrew C and Rohl A. 2001. Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Mol Biol Evol*. 18: 1864–1881.

Friedlaender J, Schurr T, Gentz F, Koki G, Friedlaender F, Horvat G, Babb P, Cerchio S, Kaestle F, Schanfield M, et al. 2005. Expanding Southwest Pacific mitochondrial haplogroups P and Q. *Mol Biol Evol*. 22: 1506–1517.

Friedlaender JS, Friedlaender FR, Hodgson JA, Stoltz M, Koki G, Horvat G, Zhadanov S, Schurr TG and Merriwether DA. 2007. Melanesian mtDNA complexity. *PLoS ONE*. 2: e248.

Frost P. 2007. Human skin-color sexual dimorphism: a test of the sexual selection hypothesis. *Am J Phys Anthropol*. 133: 779–780; author reply 780–771.

Fuller BB, Spaulding DT and Smith DR. 2001. Regulation of the catalytic activity of preexisting tyrosinase in black and Caucasian human melanocyte cell cultures. *Exp Cell Res*. 262: 197–208.

Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admetlla A, Pattini L and Nielsen R. 2011. Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genet*. 7: e1002355.

Gerstenblith MR, Goldstein AM, Fargnoli MC, Peris K and Landi MT. 2007. Comprehensive evaluation of allele frequency differences of MC1R variants across populations. *Hum Mutat*. 28: 495–505.

Giles RE, Blanc H, Cann HM and Wallace DC. 1980. Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci U S A*. 77: 6715–6719.

Ginger RS, Askew SE, Ogborne RM, Wilson S, Ferdinando D, Dadd T, Smith AM, Kazi S, Szerencsei RT, Winkfein RJ, et al. 2008. SLC24A5 encodes a trans-Golgi network protein with potassium-dependent sodium-calcium exchange activity that regulates human epidermal melanogenesis. *J Biol Chem*. 283: 5486–5495.

Glover IC and Presland G. 1985. Microliths in Indonesian flaked stone industries. *Recent Advances in Indo-Pacific Prehistory*. Misra VN and Bellwoood P. Oxford & IBH, New Delhi: 185–195.

Gollan K. 1985. Prehistoric dogs in Australia: an Indian Origin? *Recent Advances in Indo-Pacific Prehistory*. Misra VN and Bellwoood P. Oxford and IBH, New Delhi: 439–443.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, et al. 2010. A draft sequence of the Neandertal genome. *Science*. 328: 710–722.

Grossman SR, Andersen KG, Shlyakhter I, Tabrizi S, Winnicki S, Yen A, Park DJ, Griesemer D, Karlsson EK, Wong SH, et al. 2013. Identifying recent adaptations in large-scale genomic data. *Cell*. 152: 703–713.

Hammer MF. 1994. A Recent Insertion of an Alu Element on the Y-Chromosome Is a Useful Marker for Human-Population Studies. *Mol Biol Evol*. 11: 749–761.

Hancock AM, Witonsky DB, Ehler E, Alkorta-Aranburu G, Beall C, Gebremedhin A, Sukernik R, Utermann G, Pritchard J, Coop G, et al. 2010. Colloquium paper: human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proc Natl Acad Sci U S A*. 107 Suppl 2: 8924–8930.

Hancock AM, Witonsky DB, Gordon AS, Eshel G, Pritchard JK, Coop G and Di Rienzo A. 2008. Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet*. 4: e32.

Harding RM, Healy E, Ray AJ, Ellis NS, Flanagan N, Todd C, Dixon C, Sajantila A, Jackson IJ, Birch-Machin MA, et al. 2000. Evidence for variable selective pressures at MC1R. *Am J Hum Genet*. 66: 1351–1361.

Henn BM, Cavalli-Sforza LL and Feldman MW. 2012. The great human expansion. *Proc Natl Acad Sci U S A*. 109: 17758–17764.

Hill C, Soares P, Mormina M, Macaulay V, Clarke D, Blumbach PB, Vizuete-Forster M, Forster P, Bulbeck D, Oppenheimer S, et al. 2007. A mitochondrial stratigraphy for island southeast Asia. *Am J Hum Genet*. 80: 29–43.

Hiscock P. 2008. *The archaeology of ancient Australia*. London, New York, Routledge.

Holick MF and Chen TC. 2008. Vitamin D deficiency: a worldwide problem with health consequences. *Am J Clin Nutr*. 87: 1080S-1086S.

Huoponen K, Schurr TG, Chen Y and Wallace DC. 2001. Mitochondrial DNA variation in an aboriginal Australian population: evidence for genetic isolation and regional differentiation. *Hum Immunol*. 62: 954–969.

Huxley TH. 1870. On the Geographical Distribution of the Chief Modifications of Mankind. *Journal of the Ethnological Society of London*. 2: 404–412.

Ingman M and Gyllensten U. 2003. Mitochondrial genome variation and evolutionary history of Australian and New Guinean aborigines. *Genome Res*. 13: 1600–1606.

Ingman M, Kaessmann H, Pääbo S and Gyllensten U. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature*. 408: 708–713.

Ito S and Wakamatsu K. 2003. Quantitative Analysis of Eumelanin and Pheomelanin in Humans, Mice, and Other Animals: a Comparative Review. *Pigment Cell Res*. 16: 523–531.

Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW and Parsons TJ. 1996. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. *Nat Genet*. 12: 417–420.

Izagirre N, Garcia I, Junquera C, de la Rua C and Alonso S. 2006. A scan for signatures of positive selection in candidate loci for skin pigmentation in humans. *Mol Biol Evol*. 23: 1697–1706.

Jablonski NG. 2004. The Evolution of Human Skin and Skin Color. *Annual Review of Anthropology*. 33: 585–623.

Jablonski NG and Chaplin G. 2010. Colloquium paper: human skin pigmentation as an adaptation to UV radiation. *Proc Natl Acad Sci U S A*. 107 Suppl 2: 8962–8968.

Jarvis JP, Scheinfeldt LB, Soi S, Lambert C, Omberg L, Ferwerda B, Froment A, Bodo JM, Beggs W, Hoffman G, et al. 2012. Patterns of ancestry, signatures of natural selection, and genetic association with stature in Western African pygmies. *PLoS Genet*. 8: e1002641.

Jinam TA, Hong LC, Phipps ME, Stoneking M, Ameen M, Edo J and Saitou N. 2012. Evolutionary history of continental southeast Asians: "early train" hypothesis based on genetic analysis of mitochondrial and autosomal DNA data. *Mol Biol Evol*. 29: 3513–3527.

Jobling MA, Samara V, Pandya A, Fretwell N, Bernasconi B, Mitchell RJ, Gerelsaikhan T, Dashnyam B, Sajantila A, Salo PJ, et al. 1996. Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. *Hum Mol Genet*. 5: 1767–1775.

Jobling MA and Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet*. 4: 598–612.

Juzeniene A, Setlow R, Porojnicu A, Steindal AH and Moan J. 2009. Development of different human skin colors: a review highlighting photobiological and photobiophysical aspects. *J Photochem Photobiol B*. 96: 93–100.

Juzeniene A, Stokke KT, Thune P and Moan J. 2010. Pilot study of folate status in healthy volunteers and in patients with psoriasis before and after UV exposure. *J Photochem Photobiol B*. 101: 111–116.

Karafet TM, Hallmark B, Cox MP, Sudoyo H, Downey S, Lansing JS and Hammer MF. 2010. Major east-west division underlies Y chromosome stratification across Indonesia. *Mol Biol Evol*. 27: 1833–1844.

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL and Hammer MF. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res*. 18: 830–838.

Kayser M. 2010. The human genetic history of Oceania: near and remote views of dispersal. *Curr Biol*. 20: R194–201.

Kayser M, Brauer S, Weiss G, Schiefenhovel W, Underhill PA and Stoneking M. 2001. Independent histories of human Y chromosomes from Melanesia and Australia. *Am J Hum Genet*. 68: 173–190.

Kayser M, Choi Y, van Oven M, Mona S, Brauer S, Trent RJ, Suarkia D, Schiefenhovel W and Stoneking M. 2008. The impact of the Austronesian expansion: evidence from mtDNA and Y chromosome diversity in the Admiralty Islands of Melanesia. *Mol Biol Evol*. 25: 1362–1374.

Khan R and Khan BS. 2010. Diet, disease and pigment variation in humans. *Med Hypotheses*. 75: 363–367.

Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K, Passarino G, Underhill PA, Scharfe C, Torroni A, et al. 2006. The role of selection in the evolution of human mitochondrial genomes. *Genetics*. 172: 373–387.

Kong QP, Sun C, Wang HW, Zhao M, Wang WZ, Zhong L, Hao XD, Pan H, Wang SY, Cheng YT, et al. 2011. Large-scale mtDNA screening reveals a surprising matrilineal complexity in east Asia and its implications to the peopling of the region. *Mol Biol Evol*. 28: 513–522.

Kong QP, Yao YG, Sun C, Bandelt HJ, Zhu CL and Zhang YP. 2003. Phylogeny of east Asian mitochondrial DNA lineages inferred from complete sequences. *Am J Hum Genet*. 73: 671–676.

Kumar S, Ravuri RR, Koneru P, Urade BP, Sarkar BN, Chandrasekar A and Rao VR. 2009. Reconstructing Indian-Australian phylogenetic link. *BMC Evol Biol*. 9: 173.

Lalueza-Fox C, Rompler H, Caramelli D, Staubert C, Catalano G, Hughes D, Rohland N, Pilli E, Longo L, Condemi S, et al. 2007. A melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. *Science*. 318: 1453–1455.

Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, Aros MC, Jurynec MJ, Mao X, Humphreville VR, Humbert JE, et al. 2005. SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science*. 310: 1782–1786.

Lao O, de Gruijter JM, van Duijn K, Navarro A and Kayser M. 2007. Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. *Ann Hum Genet*. 71: 354–369.

Liu PT, Stenger S, Li H, Wenzel L, Tan BH, Krutzik SR, Ochoa MT, Schauber J, Wu K, Meinken C, et al. 2006. Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response. *Science*. 311: 1770–1773.

Loogväli EL, Kivisild T, Margus T and Villems R. 2009. Explaining the imperfection of the molecular clock of hominid mitochondria. *PLoS ONE*. 4: e8260.

Loomis WF. 1967. Skin-pigment regulation of vitamin-D biosynthesis in man. *Science*. 157: 501–506.

Lu D, Willard D, Patel IR, Kadwell S, Overton L, Kost T, Luther M, Chen W, Woychik RP, Wilkison WO, et al. 1994. Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. *Nature*. 371: 799–802.

Lucotte G, Mercier G, Dieterlen F and Yuasa I. 2010. A decreasing gradient of 374F allele frequencies in the skin pigmentation gene SLC45A2, from the north of West Europe to North Africa. *Biochem Genet*. 48: 26–33.

Lum JK and Cann RL. 2000. mtDNA lineage analyses: origins and migrations of Micronesians and Polynesians. *Am J Phys Anthropol*. 113: 151–168.

Lum JK, Rickards O, Ching C and Cann RL. 1994. Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. *Hum Biol*. 66: 567–590.

Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, Blackburn J, Semino O, Scozzari R, Cruciani F, et al. 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science*. 308: 1034–1036.

Madrigal L and Kelly W. 2007. Human skin-color sexual dimorphism: a test of the sexual selection hypothesis. *Am J Phys Anthropol*. 132: 470–482.

Madrigal L and Kelly W. 2007. Human skin-color sexual dimorphism: a test of the sexual selection hypothesis. Reply to Frost (2007). *Am J Phys Anthropol*. 133: 780–781.

Majumder PP. 2010. The human genetic history of South Asia. *Curr Biol*. 20: R184–187.

Makova K and Norton H. 2005. Worldwide polymorphism at the MC1R locus and normal pigmentation variation in humans. *Peptides*. 26: 1901–1908.

McEvoy B, Beleza S and Shriver MD. 2006. The genetic architecture of normal variation in human pigmentation: an evolutionary perspective and model. *Hum Mol Genet*. 15 Spec No 2: R176–181.

McEvoy BP, Lind JM, Wang ET, Moyzis RK, Visscher PM, van Holst Pellekaan SM and Wilton AN. 2010. Whole-genome genetic diversity in a sample of Australians with deep Aboriginal ancestry. *Am J Hum Genet*. 87: 297–305.

Mendez FL, Krahn T, Schrack B, Krahn AM, Veeramah KR, Woerner AE, Fomine FL, Bradman N, Thomas MG, Karafet TM, et al. 2013. An African American paternal lineage adds an extremely ancient root to the human Y chromosome phylogenetic tree. *Am J Hum Genet*. 92: 454–459.

Merriwether DA, Hodgson JA, Friedlaender FR, Allaby R, Cerchio S, Koki G and Friedlaender JS. 2005. Ancient mitochondrial M haplogroups identified in the Southwest Pacific. *Proc Natl Acad Sci U S A*. 102: 13034–13039.

Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, Kaldma K, Serk P, Karmin M, Behar DM, Gilbert MT, et al. 2004. Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet*. 5: 26.

Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Prufer K, de Filippo C, et al. 2012. A high-coverage genome sequence from an archaic Denisovan individual. *Science*. 338: 222–226.

Migliano AB, Gallego Romero I, Metspalu M, Leavesley M, Pagani L, Antao T, Huang W, Sherman BT, Siddle K, Scholes C, et al. in press. Evolution of the pygmy phenotype: evidence of positive selection from genome-wide scans in African, Asian and Melanesian pygmies. *Hum Biol*.

Mijares AS, Detroit F, Piper P, Grun R, Bellwood P, Aubert M, Champion G, Cuevas N, De Leon A and Dizon E. 2010. New evidence for a 67,000-year-old human presence at Callao Cave, Luzon, Philippines. *J Hum Evol*. 59: 123–132.

Miller CT, Beleza S, Pollen AA, Schluter D, Kittles RA, Shriver MD and Kingsley DM. 2007. cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell*. 131: 1179–1189.

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, Easley K, Chen E, Brown MD, et al. 2003. Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci U S A*. 100: 171–176.

Mona S, Grunz KE, Brauer S, Pakendorf B, Castri L, Sudoyo H, Marzuki S, Barnes RH, Schmidtke J, Stoneking M, et al. 2009. Genetic admixture history of Eastern Indonesia as revealed by Y-chromosome and mitochondrial DNA analysis. *Mol Biol Evol*. 26: 1865–1877.

Montoliu L, Oetting WS and Bennett DC. 2012. *Color Genes*. Retrieved March 25, 2013, from http://www.espcr.org/micemut.

Moodley Y, Linz B, Yamaoka Y, Windsor HM, Breurec S, Wu JY, Maady A, Bernhoft S, Thiberge JM, Phuanukoonnon S, et al. 2009. The peopling of the Pacific from a bacterial perspective. *Science*. 323: 527–530.

Mulvaney DJ and Kamminga J. 1999. *Prehistory of Australia*. Washington, D.C., Smithsonian Institution Press.

Myles S, Somel M, Tang K, Kelso J and Stoneking M. 2007. Identifying genes underlying skin pigmentation differences among human populations. *Hum Genet*. 120: 613–621.

National Institutes of Health. 2013. *Dietary Supplement Fact Sheet: Folate*. Retrieved March 28, 2013, from http://ods.od.nih.gov/factsheets/Folate-HealthProfessional/.

Norton HL and Hammer MF. 2007. Sequence variation in the pigmentation candidate gene SLC24A5 and evidence for independent evolution of light skin in European and East Asian populations. *77th Annual Meeting of the American Association of Physical Anthropologists*.

Norton HL, Kittles RA, Parra E, McKeigue P, Mao X, Cheng K, Canfield VA, Bradley DG, McEvoy B and Shriver MD. 2007. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Mol Biol Evol*. 24: 710–722.

O'Connell JF and Allen J. 2004. Dating the colonization of Sahul (Pleistocene Australia–New Guinea): a review of recent research. *Journal of Archaeological Science*. 31: 835–853.

Oppenheimer S. 2012. A single southern exit of modern humans from Africa: Before or after Toba? *Quaternary International*. 258: 88–99.

Ortonne JP. 2002. Photoprotective properties of skin melanin. *Br J Dermatol*. 146 Suppl 61: 7–10.

Oskarsson MC, Klutsch CF, Boonyaprakob U, Wilton A, Tanabe Y and Savolainen P. 2012. Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. *Proc Biol Sci*. 279: 967–974.

Palanichamy MG, Sun C, Agrawal S, Bandelt HJ, Kong QP, Khan F, Wang CY, Chaudhuri TK, Palla V and Zhang YP. 2004. Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. *Am J Hum Genet*. 75: 966–978.

Parra EJ. 2007. Human pigmentation variation: evolution, genetic basis, and implications for public health. *Am J Phys Anthropol*. Suppl 45: 85–105.

Paul LM, Simons GF and Fennig CD. 2013. *Ethnologue: Languages of the World*. Dallas, Texas, SIL International.

Payne BA, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, Taylor RW, Samuels DC, Santibanez-Koref M and Chinnery PF. 2013. Universal heteroplasmy of human mitochondrial DNA. *Hum Mol Genet*. 22: 384–390.

Peng MS, Quang HH, Dang KP, Trieu AV, Wang HW, Yao YG, Kong QP and Zhang YP. 2010. Tracing the Austronesian footprint in Mainland Southeast Asia: a perspective from mitochondrial DNA. *Mol Biol Evol*. 27: 2417–2430.

Pickrell JK, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, Srinivasan BS, Barsh GS, Myers RM, Feldman MW, et al. 2009. Signals of recent positive selection in a worldwide sample of human populations. *Genome Res*. 19: 826–837.

Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, Ruczinski I, Beaty TH, Mathias R, Reich D and Myers S. 2009. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet*. 5: e1000519.

Pugach I, Delfin F, Gunnarsdottir E, Kayser M and Stoneking M. 2013. Genome-wide data substantiate Holocene gene flow from India to Australia. *Proc Natl Acad Sci U S A*. 110: 1803–1808.

Rana B, Hewett-Emmett D, Jin L, Chang B, Sambuughin N, Lin M, Watkins S, Bamshad M, Jorde L, Ramsay M, et al. 1999. High polymorphism at the human melanocortin 1 receptor locus. *Genetics*. 151: 1547–1557.

Rasmussen M, Guo X, Wang Y, Lohmueller KE, Rasmussen S, Albrechtsen A, Skotte L, Lindgreen S, Metspalu M, Jombart T, et al. 2011. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science*. 334: 94–98.

Ray K, Chaki M and Sengupta M. 2007. Tyrosinase and ocular diseases: some novel thoughts on the molecular basis of oculocutaneous albinism type 1. *Prog Retin Eye Res*. 26: 323–358.

Redd AJ, Roberts-Thomson J, Karafet T, Bamshad M, Jorde LB, Naidu JM, Walsh B and Hammer MF. 2002. Gene flow from the Indian subcontinent to Australia: evidence from the Y chromosome. *Curr Biol*. 12: 673–677.

Redd AJ and Stoneking M. 1999. Peopling of Sahul: mtDNA variation in aboriginal Australian and Papua New Guinean populations. *Am J Hum Genet*. 65: 808–828.

Rees JL and Harding RM. 2012. Understanding the evolution of human pigmentation: recent contributions from population genetics. *J Invest Dermatol*. 132: 846–853.

Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PL, et al. 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*. 468: 1053–1060.

Reich D, Patterson N, Kircher M, Delfin F, Nandineni MR, Pugach I, Ko AM, Ko YC, Jinam TA, Phipps ME, et al. 2011. Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am J Hum Genet*. 89: 516–528.

Relethford JH. 2002. Apportionment of global human genetic diversity based on craniometrics and skin color. *Am J Phys Anthropol*. 118: 393–398.

Roberts-Thomson J, Martinson J, Norwich J, Harding R, Clegg J and Boettcher B. 1996. An ancient common origin of aboriginal Australians and New Guinea highlanders is supported by alpha-globin haplotype analysis. *Am J Hum Genet*. 58: 1017–1024.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, et al. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature*. 449: 913–918.

Savage SA, Gerstenblith MR, Goldstein AM, Mirabello L, Fargnoli MC, Peris K and Landi MT. 2008. Nucleotide diversity and population differentiation of the melanocortin 1 receptor gene, MC1R. *BMC Genet*. 9: 31.

Savolainen P, Leitner T, Wilton AN, Matisoo-Smith E and Lundeberg J. 2004. A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. *Proc Natl Acad Sci U S A*. 101: 12387–12390.

Sayres M. 2013. Timing of ancient human Y lineage depends on the mutation rate: A comment on Mendez et al. *arXiv preprint arXiv:1304.6098*.

Scally A and Durbin R. 2012. Revising the human mutation rate: implications for understanding human evolution. *Nat Rev Genet*. 13: 745–753.

Schaffner SF, Foo C, Gabriel S, Reich D, Daly MJ and Altshuler D. 2005. Calibrating a coalescent simulation of human genome sequence variation. *Genome Res*. 15: 1576–1583.

Scherer D and Kumar R. 2010. Genetics of pigmentation in skin cancer – a review. *Mutat Res*. 705: 141–153.

Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, Chow CE, Lin AA, Mitra M, Sil SK, Ramesh A, et al. 2006. 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*. 78: 202–221.

Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, et al. 2003. Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum Genet*. 112: 387–399.

Simonson TS, Yang Y, Huff CD, Yun H, Qin G, Witherspoon DJ, Bai Z, Lorenzo FR, Xing J, Jorde LB, et al. 2010. Genetic evidence for high-altitude adaptation in Tibet. *Science*. 329: 72–75.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*. 423: 825–837.

Soares P, Alshamali F, Pereira JB, Fernandes V, Silva NM, Afonso C, Costa MD, Musilova E, Macaulay V, Richards MB, et al. 2012. The Expansion of mtDNA Haplogroup L3 within and out of Africa. *Mol Biol Evol*. 29: 915–927.

Soares P, Ermini L, Thomson N, Mormina M, Rito T, Rohl A, Salas A, Oppenheimer S, Macaulay V and Richards MB. 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet*. 84: 740–759.

Soares P, Rito T, Trejaut J, Mormina M, Hill C, Tinkler-Hundal E, Braid M, Clarke DJ, Loo JH, Thomson N, et al. 2011. Ancient voyaging and Polynesian origins. *Am J Hum Genet*. 88: 239–247.

Soejima M, Tachida H, Ishida T, Sano A and Koda Y. 2006. Evidence for recent positive selection at the human AIM1 locus in a European population. *Mol Biol Evol*. 23: 179–188.

Specht J and Gosden C. 1997. Dating Lapita Pottery in the Bismark Archipelago, Papua New Guinea. *Asian Perspectives*. 36: 175–199.

St John JC, Facucho-Oliveira J, Jiang Y, Kelly R and Salah R. 2010. Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. *Hum Reprod Update*. 16: 488–509.

Steindal AH and Moan J. 2008. The evolution of different skin colours. *Solar Radiation and Human Health*, Oslo, Norway, The Norwegian Academy of Science and Letters.

Stokowski RP, Pant PV, Dadd T, Fereday A, Hinds DA, Jarman C, Filsell W, Ginger RS, Green MR, van der Ouderaa FJ, et al. 2007. A genomewide association study of skin pigmentation in a South Asian population. *Am J Hum Genet*. 81: 1119–1132.

Stoneking M, Jorde LB, Bhatia K and Wilson AC. 1990. Geographic variation in human mitochondrial DNA from Papua New Guinea. *Genetics*. 124: 717–733.

Stringer C. 2000. Palaeoanthropology. Coasting out of Africa. *Nature*. 405: 24–25, 27.

Sturm RA. 2009. Molecular genetics of human pigmentation diversity. *Hum Mol Genet*. 18: R9–17.

Sturm RA and Duffy DL. 2012. Human pigmentation genes under environmental selection. *Genome Biol*. 13: 248.

Sturm RA, Duffy DL, Zhao ZZ, Leite FP, Stark MS, Hayward NK, Martin NG and Montgomery GW. 2008. A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. *Am J Hum Genet*. 82: 424–431.

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G, et al. 2007. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet*. 39: 1443–1452.

Summerhayes GR, Leavesley M, Fairbairn A, Mandui H, Field J, Ford A and Fullagar R. 2010. Human adaptation and plant use in highland New Guinea 49,000 to 44,000 years ago. *Science*. 330: 78–81.

Sykes B, Leiboff A, Low-Beer J, Tetzner S and Richards M. 1995. The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. *Am J Hum Genet*. 57: 1463–1475.

Tabbada KA, Trejaut J, Loo JH, Chen YM, Lin M, Mirazon-Lahr M, Kivisild T and De Ungria MC. 2010. Philippine mitochondrial DNA diversity: a populated viaduct between Taiwan and Indonesia? *Mol Biol Evol*. 27: 21–31.

Taylor D, Nagle N, Ballantyne KN, van Oorschot RA, Wilcox S, Henry J, Turakulov R and Mitchell RJ. 2012. An investigation of admixture in an Australian Aboriginal Y-chromosome STR database. *Forensic Sci Int Genet*. 6: 532–538.

Tommaseo-Ponzetta M, Attimonelli M, De Robertis M, Tanzariello F and Saccone C. 2002. Mitochondrial DNA variability of West New Guinea populations. *Am J Phys Anthropol*. 117: 49–67.

Torroni A, Achilli A, Macaulay V, Richards M and Bandelt HJ. 2006. Harvesting the fruit of the human mtDNA tree. *Trends Genet*. 22: 339–345.

Trejaut JA, Kivisild T, Loo JH, Lee CL, He CL, Hsu CJ, Lee ZY and Lin M. 2005. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biol*. 3: e247.

Tripathi RK, Hearing VJ, Urabe K, Aroca P and Spritz RA. 1992. Mutational mapping of the catalytic activities of human tyrosinase. *J Biol Chem*. 267: 23707–23712.

Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL and Oefner PJ. 1997. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res*. 7: 996–1005.

Underhill PA and Kivisild T. 2007. Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet*. 41: 539–564.

Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ and Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet*. 65: 43–62.

UniProt Consortium. 2013. Retrieved March 30, 2013, from http://www.uniprot.org.

Valverde P, Healy E, Jackson I, Rees JL and Thody AJ. 1995. Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nat Genet*. 11: 328–330.

van Holst Pellekaan S. 2013. Genetic evidence for the colonization of Australia. *Quaternary International*. 285: 44–56.

van Holst Pellekaan S, Frommer M, Sved J and Boettcher B. 1998. Mitochondrial control-region sequence variation in aboriginal Australians. *Am J Hum Genet*. 62: 435–449.

van Holst Pellekaan SM. 2012. Socially responsible genetic research with descendants of the First Australians. *Investig Genet*. 3: 22.

van Holst Pellekaan SM, Ingman M, Roberts-Thomson J and Harding RM. 2006. Mitochondrial genomics identifies major haplogroups in Aboriginal Australians. *Am J Phys Anthropol*. 131: 282–294.

van Oven M and Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 30: E386–394.

Vigilant L, Stoneking M, Harpending H, Hawkes K and Wilson AC. 1991. African populations and the evolution of human mitochondrial DNA. *Science*. 253: 1503–1507.

Voight BF, Kudaravalli S, Wen X and Pritchard JK. 2006. A map of recent positive selection in the human genome. *PLoS Biol*. 4: e72.

Wei W, Ayub Q, Chen Y, McCarthy S, Hou Y, Carbone I, Xue Y and Tyler-Smith C. 2013. A calibrated human Y-chromosomal phylogeny based on resequencing. *Genome Res*. 23: 388–395.

Williamson SH, Hubisz MJ, Clark AG, Payseur BA, Bustamante CD and Nielsen R. 2007. Localizing recent adaptive evolution in the human genome. *PLoS Genet*. 3: e90.

Wong LP, Ong RT, Poh WT, Liu X, Chen P, Li R, Lam KK, Pillai NE, Sim KS, Xu H, et al. 2013. Deep whole-genome sequencing of 100 southeast Asian Malays. *Am J Hum Genet*. 92: 52–66.

Xue Y, Wang Q, Long Q, Ng BL, Swerdlow H, Burton J, Skuce C, Taylor R, Abdellah Z, Zhao Y, et al. 2009. Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Curr Biol*. 19: 1453–1457.

Yuasa I, Harihara S, Jin F, Nishimukai H, Fujihara J, Fukumori Y, Takeshita H, Umetsu K and Saitou N. 2011. Distribution of OCA2*481Thr and OCA2*615Arg, associated with hypopigmentation, in several additional populations. *Leg Med (Tokyo)*. 13: 215–217.

Yuen AW and Jablonski NG. 2010. Vitamin D: in the evolution of human skin colour. *Med Hypotheses*. 74: 39–44.

Yunusbayev B, Metspalu M, Jarve M, Kutuev I, Rootsi S, Metspalu E, Behar DM, Varendi K, Sahakyan H, Khusainova R, et al. 2012. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol*. 29: 359–365.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors Prof. Toomas Kivisild and Prof. Richard Villems for professional guidance, fruitful and critical discussions and patience. It has been a great pleasure to work together on this interesting topic. I would like to give special thanks to Kristiina Tambets, Mari Järve, Monika Karmin and Mait Metspalu (*aitäh!*), and Erwan Pennarun (*merci!*) for their kind support and help, energy, and positive attitude. This work would have been impossible to conduct without the technical assistance of Ille Hilpus, Tuuli Reisberg and Viljo Soo. I would also like to thank Reedik Mägi, Irene Gallego Romero and the High Performance Computing Center in the University of Tartu for help with software programming and analytical pipelines. I am very thankful to my first supervisor Jüri Parik for his wet-lab advice and, of course, the impressive photographic skills and equipment. I am very grateful to my Indian colleagues, Chandana Basu Mallick and Gyaneshwer Chaubey, for providing South Asian samples and help with South Asian human maternal and paternal phylogeny. I owe many thanks to all my co-authors and colleagues, including members of the Evolutionary Biology group: Ene Metspalu, Siiri Rootsi, Maere Reidla, Eva-Liis Loogväli, Erika Tamm, Anu Solnik and Anne-Mai Ilumäe.

Огромное спасибо моим русскоговорящим коллегам, Алёне и Баязиту, за хорошее настроение и возможность для (около)научного и неформального диалога на родном для меня языке. Я также отдельно благодарен профессору Рихарду Виллемсу за его прекрасное многоязычное общение.

Finally, yet importantly, I would like to express my heartfelt thanks to my beloved parents, brother, my wife and daughter for all the love and support. Спасибо! Эта работа была бы невозможна без вашей поддержки, любви и терпения!

PUBLICATIONS

CURRICULUM VITAE

Name: Georgi Hudjašov
Date of birth: November 25, 1980
Address: Department of Evolutionary Biology
Institute of Molecular and Cell Biology
University of Tartu and the Estonian Biocentre
Riia 23, 51010 Tartu, Estonia
Phone: +372 737 5054
E-mail: hudja@ebc.ee

Academic background:

Since 2006	University of Tartu, Faculty of Science and Technology, Molecular and Cell Biology, PhD student
2004–2006	University of Tartu, Faculty of Biology and Geography, Gene Technology, MSc
1998–2003	University of Tartu, Faculty of Biology and Geography, Biology, BSc

Distinctions and awards:

2006, "Kaleva Travel" travel grant
2006, Estonian Academy of Sciences scholarship
2005, Artur Lind scholarship, Estonian Genome Foundation

Publications:

Mallick CB, Iliescu FM, Möls M, Hill S, Tamang R, Chaubey G, Goto R, Ho SY, Romero IG, Crivellaro F, **Hudjashov G**, Rai N, Metspalu M, Mascie-Taylor N, Pitchappan R, Singh L, Lahr MM, Thangaraj K, Villems R and Kivisild T. The Light Skin Allele of *SLC24A5* in South Asians and Europeans Shares Identity by Descent. *PLoS Genet*, accepted for publication.

Migliano AB, Gallego Romero I, Metspalu M, Leavesley M, Pagani L, Antao T, Huang W, Sherman BT, Siddle K, Scholes C, **Hudjashov G**, Kaitokai E, Mandui H, Babalu A, Belatti M, Cagan A, Hopkinshaw B, Shaw C, Tabbada KA, Nelis M, Metspalu E, Mägi R, Lempicki RA, Villems R, Lahr MM and Kivisild T. Evolution of the pygmy phenotype: evidence of positive selection from genome-wide scans in African, Asian and Melanesian pygmies. *Hum Biol*, accepted for publication.

Hudjashov G, Villems R and Kivisild T. 2013. Global patterns of diversity and selection in human tyrosinase gene. *PLoS ONE*. 8: e74307.

Chaubey G, Metspalu M, Choi Y, Magi R, Romero IG, Soares P, van Oven M, Behar DM, Rootsi S, **Hudjashov G**, Mallick CB, Karmin M, Nelis M, Parik J, Reddy AG, Metspalu E, van Driem G, Xue Y, Tyler-Smith C, Thangaraj K, Singh L, Remm M, Richards MB, Lahr MM, Kayser M, Villems R and Kivisild T. 2011. Population genetic structure in Indian Austroasiatic

speakers: the role of landscape barriers and sex-specific admixture. *Mol Biol Evol*. 28: 1013–1024.

Metspalu M, Romero IG, Yunusbayev B, Chaubey G, Mallick CB, **Hudjashov G**, Nelis M, Magi R, Metspalu E, Remm M, Pitchappan R, Singh L, Thangaraj K, Villems R and Kivisild T. 2011. Shared and unique components of human population structure and genome-wide signals of positive selection in South Asia. *Am J Hum Genet*. 89: 731–744.

Soares P, Trejaut JA, Loo JH, Hill C, Mormina M, Lee CL, Chen YM, **Hudjashov G**, Forster P, Macaulay V, Bulbeck D, Oppenheimer S, Lin M and Richards MB. 2008. Climate change and postglacial human dispersals in southeast Asia. *Mol Biol Evol*. 25: 1209–1218.

Hudjashov G, Kivisild T, Underhill PA, Endicott P, Sanchez JJ, Lin AA, Shen P, Oefner P, Renfrew C, Villems R and Forster P. 2007. Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. *Proc Natl Acad Sci U S A*. 104: 8726–8730.

Metspalu M, Kivisild T, Metspalu E, Parik J, **Hudjashov G**, Kaldma K, Serk P, Karmin M, Behar DM, Gilbert MT, Endicott P, Mastana S, Papiha SS, Skorecki K, Torroni A and Villems R. 2004. Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet*. 5: 26.

ELULOOKIRJELDUS

Nimi: Georgi Hudjašov
Sünniaeg: 25. november, 1980
Aadress: Evolutsioonilise bioloogia õppetool
Molekulaar- ja Rakubioloogia Instituut
Tartu Ülikool ja Eesti Biokeskus
Riia 23, 51010 Tartu, Eesti
Telefon: +372 737 5054
E-post: hudja@ebc.ee

Akadeemiline elukäik:

Alates 2006	Tartu Ülikool, Loodus- ja tehnoloogiateaduskond, Molekulaar- ja Rakubioloogia, doktorant
2004–2006	Tartu Ülikool, Bioloogia-geograafiateaduskond, Geenitehnoloogia, MSc
1998–2003	Tartu Ülikool, Bioloogia-geograafiateaduskond, Bioloogia, BSc

Tunnustused ja auhinnad:

2006, Kaleva Travel reisistipeendium
2006, Eesti Teaduste Akadeemia üliõpilastööde konkurss, I koht
2005, Artur Linnu Geenitehnoloogia Stipendium, SA Geenikeskus

Publikatsioonid:

Mallick CB, Iliescu FM, Möls M, Hill S, Tamang R, Chaubey G, Goto R, Ho SY, Romero IG, Crivellaro F, **Hudjashov G**, Rai N, Metspalu M, Mascie-Taylor N, Pitchappan R, Singh L, Lahr MM, Thangaraj K, Villems R ja Kivisild T. The Light Skin Allele of *SLC24A5* in South Asians and Europeans Shares Identity by Descent. *PLoS Genet*, avaldamiseks vastu võetud.

Migliano AB, Gallego Romero I, Metspalu M, Leavesley M, Pagani L, Antao T, Huang W, Sherman BT, Siddle K, Scholes C, **Hudjashov G**, Kaitokai E, Mandui H, Babalu A, Belatti M, Cagan A, Hopkinshaw B, Shaw C, Tabbada KA, Nelis M, Metspalu E, Mägi R, Lempicki RA, Villems R, Lahr MM ja Kivisild T. Evolution of the pygmy phenotype: evidence of positive selection from genome-wide scans in African, Asian and Melanesian pygmies. *Hum Biol*, avaldamiseks vastu võetud.

Hudjashov G, Villems R ja Kivisild T. 2013. Global patterns of diversity and selection in human tyrosinase gene. *PLoS ONE*. 8: e74307.

Chaubey G, Metspalu M, Choi Y, Magi R, Romero IG, Soares P, van Oven M, Behar DM, Rootsi S, **Hudjashov G**, Mallick CB, Karmin M, Nelis M, Parik J, Reddy AG, Metspalu E, van Driem G, Xue Y, Tyler-Smith C, Thangaraj K, Singh L, Remm M, Richards MB, Lahr MM, Kayser M, Villems R ja Kivisild T. 2011. Population genetic structure in Indian Austroasiatic speakers: the role of landscape barriers and sex-specific admixture. *Mol Biol Evol*. 28: 1013–1024.

Metspalu M, Romero IG, Yunusbayev B, Chaubey G, Mallick CB, **Hudjashov G**, Nelis M, Magi R, Metspalu E, Remm M, Pitchappan R, Singh L, Thangaraj K, Villems R ja Kivisild T. 2011. Shared and unique components of human population structure and genome-wide signals of positive selection in South Asia. *Am J Hum Genet*. 89: 731–744.

Soares P, Trejaut JA, Loo JH, Hill C, Mormina M, Lee CL, Chen YM, **Hudjashov G**, Forster P, Macaulay V, Bulbeck D, Oppenheimer S, Lin M ja Richards MB. 2008. Climate change and postglacial human dispersals in southeast Asia. *Mol Biol Evol*. 25: 1209–1218.

Hudjashov G, Kivisild T, Underhill PA, Endicott P, Sanchez JJ, Lin AA, Shen P, Oefner P, Renfrew C, Villems R ja Forster P. 2007. Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. *Proc Natl Acad Sci U S A*. 104: 8726–8730.

Metspalu M, Kivisild T, Metspalu E, Parik J, **Hudjashov G**, Kaldma K, Serk P, Karmin M, Behar DM, Gilbert MT, Endicott P, Mastana S, Papiha SS, Skorecki K, Torroni A ja Villems R. 2004. Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet*. 5: 26.

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand.** The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro.* Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous grassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.
41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.

42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.
61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.

63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p
82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.
83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.

84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.
85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.
86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.
87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.
88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.
89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004. 117 p.
90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004. 175 p.
91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004. 109 p.
92. **Kristiina Tambets.** Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004. 163 p.
93. **Arvi Jõers.** Regulation of p53-dependent transcription. Tartu, 2004. 103 p.
94. **Lilian Kadaja.** Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004. 103 p.
95. **Jaak Truu.** Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004. 128 p.
96. **Maire Peters.** Natural horizontal transfer of the *pheBA* operon. Tartu, 2004. 105 p.
97. **Ülo Maiväli.** Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004. 130 p.
98. **Merit Otsus.** Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004. 103 p.
99. **Mikk Heidemaa.** Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004. 167 p.
100. **Ilmar Tõnno.** The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004. 111 p.
101. **Lauri Saks.** Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004. 144 p.
102. **Siiri Rootsi.** Human Y-chromosomal variation in European populations. Tartu, 2004. 142 p.
103. **Eve Vedler.** Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.
104. **Andres Tover.** Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 126 p.
105. **Helen Udras.** Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005. 100 p.

106. **Ave Suija.** Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005. 162 p.
107. **Piret Lõhmus.** Forest lichens and their substrata in Estonia. Tartu, 2005. 162 p.
108. **Inga Lips.** Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005. 156 p.
109. **Kaasik, Krista.** Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005. 121 p.
110. **Juhan Javoiš.** The effects of experience on host acceptance in ovipositing moths. Tartu, 2005. 112 p.
111. **Tiina Sedman.** Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005. 103 p.
112. **Ruth Aguraiuja.** Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005. 112 p.
113. **Riho Teras.** Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 106 p.
114. **Mait Metspalu.** Through the course of prehistory in india: tracing the mtDNA trail. Tartu, 2005. 138 p.
115. **Elin Lõhmussaar.** The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006. 124 p.
116. **Priit Kupper.** Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006. 126 p.
117. **Heili Ilves.** Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006. 120 p.
118. **Silja Kuusk.** Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006. 126 p.
119. **Kersti Püssa.** Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006. 90 p.
120. **Lea Tummeleht.** Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006. 94 p.
121. **Toomas Esperk.** Larval instar as a key element of insect growth schedules. Tartu, 2006. 186 p.
122. **Harri Valdmann.** Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.
123. **Priit Jõers.** Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.
124. **Kersti Lilleväli.** Gata3 and Gata2 in inner ear development. Tartu, 2007. 123 p.
125. **Kai Rünk.** Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007. 143 p.

126. **Aveliina Helm.** Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007. 89 p.
127. **Leho Tedersoo.** Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007. 233 p.
128. **Marko Mägi.** The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007. 135 p.
129. **Valeria Lulla.** Replication strategies and applications of Semliki Forest virus. Tartu, 2007. 109 p.
130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007. 79 p.
131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007. 171 p.
132. **Tatjana Krama.** Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007. 112 p.
133. **Signe Saumaa.** The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida.* Tartu, 2007. 172 p.
134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007. 96 p.
135. **Priit Kilgas.** Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007. 129 p.
136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007. 95 p.
137. **Kärt Padari.** Protein transduction mechanisms of transportans. Tartu, 2008. 128 p.
138. **Siiri-Lii Sandre.** Selective forces on larval colouration in a moth. Tartu, 2008. 125 p.
139. **Ülle Jõgar.** Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008. 99 p.
140. **Lauri Laanisto.** Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008. 133 p.
141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008. 105 p.
142. **Birgot Paavel.** Bio-optical properties of turbid lakes. Tartu, 2008. 175 p.
143. **Kaire Torn.** Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.
144. **Vladimir Vimberg.** Peptide mediated macrolide resistance. Tartu, 2008, 190 p.
145. **Daima Örd.** Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.
146. **Lauri Saag.** Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.

147. **Ulvi Karu.** Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.
148. **Jaanus Remm.** Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.
149. **Epp Moks.** Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.
150. **Eve Eensalu.** Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.
151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.
152. **Marta Putrinš.** Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.
153. **Marina Semtšenko.** Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.
154. **Marge Starast.** Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.
155. **Age Tats.** Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.
156. **Radi Tegova.** The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.
157. **Tsipe Aavik.** Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2009, 112 p.
158. **Kaja Kiiver.** Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.
159. **Meelis Kadaja.** Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.
160. **Pille Hallast.** Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.
161. **Ain Vellak.** Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.
162. **Triinu Remmel.** Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.
163. **Jaana Salujõe.** Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.
164. **Ele Vahtmäe.** Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.
165. **Liisa Metsamaa.** Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.

166. **Pille Säälik.** The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.
167. **Lauri Peil.** Ribosome assembly factors in *Escherichia coli.* Tartu, 2009, 147 p.
168. **Lea Hallik.** Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.
169. **Mariliis Tark.** Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.
170. **Riinu Rannap.** Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.
171. **Maarja Adojaan.** Molecular variation of HIV-1 and the use of this knowledge in vaccine development. Tartu, 2009, 95 p.
172. **Signe Altmäe.** Genomics and transcriptomics of human induced ovarian folliculogenesis. Tartu, 2010, 179 p.
173. **Triin Suvi.** Mycorrhizal fungi of native and introduced trees in the Seychelles Islands. Tartu, 2010, 107 p.
174. **Velda Lauringson.** Role of suspension feeding in a brackish-water coastal sea. Tartu, 2010, 123 p.
175. **Eero Talts.** Photosynthetic cyclic electron transport – measurement and variably proton-coupled mechanism. Tartu, 2010, 121 p.
176. **Mari Nelis.** Genetic structure of the Estonian population and genetic distance from other populations of European descent. Tartu, 2010, 97 p.
177. **Kaarel Krjutškov.** Arrayed Primer Extension-2 as a multiplex PCR-based method for nucleic acid variation analysis: method and applications. Tartu, 2010, 129 p.
178. **Egle Köster.** Morphological and genetical variation within species complexes: *Anthyllis vulneraria* s. l. and *Alchemilla vulgaris* (coll.). Tartu, 2010, 101 p.
179. **Erki Õunap.** Systematic studies on the subfamily Sterrhinae (Lepidoptera: Geometridae). Tartu, 2010, 111 p.
180. **Merike Jõesaar.** Diversity of key catabolic genes at degradation of phenol and *p*-cresol in pseudomonads. Tartu, 2010, 125 p.
181. **Kristjan Herkül.** Effects of physical disturbance and habitat-modifying species on sediment properties and benthic communities in the northern Baltic Sea. Tartu, 2010, 123 p.
182. **Arto Pulk.** Studies on bacterial ribosomes by chemical modification approaches. Tartu, 2010, 161 p.
183. **Maria Põllupüü.** Ecological relations of cladocerans in a brackish-water ecosystem. Tartu, 2010, 126 p.
184. **Toomas Silla.** Study of the segregation mechanism of the Bovine Papillomavirus Type 1. Tartu, 2010, 188 p.
185. **Gyaneshwer Chaubey.** The demographic history of India: A perspective based on genetic evidence. Tartu, 2010, 184 p.
186. **Katrin Kepp.** Genes involved in cardiovascular traits: detection of genetic variation in Estonian and Czech populations. Tartu, 2010, 164 p.

187. **Virve Sõber.** The role of biotic interactions in plant reproductive performance. Tartu, 2010, 92 p.
188. **Kersti Kangro.** The response of phytoplankton community to the changes in nutrient loading. Tartu, 2010, 144 p.
189. **Joachim M. Gerhold.** Replication and Recombination of mitochondrial DNA in Yeast. Tartu, 2010, 120 p.
190. **Helen Tammert.** Ecological role of physiological and phylogenetic diversity in aquatic bacterial communities. Tartu, 2010, 140 p.
191. **Elle Rajandu.** Factors determining plant and lichen species diversity and composition in Estonian *Calamagrostis* and *Hepatica* site type forests. Tartu, 2010, 123 p.
192. **Paula Ann Kivistik.** ColR-ColS signalling system and transposition of Tn*4652* in the adaptation of *Pseudomonas putida.* Tartu, 2010, 118 p.
193. **Siim Sõber.** Blood pressure genetics: from candidate genes to genome-wide association studies. Tartu, 2011, 120 p.
194. **Kalle Kipper.** Studies on the role of helix 69 of 23S rRNA in the factor-dependent stages of translation initiation, elongation, and termination. Tartu, 2011, 178 p.
195. **Triinu Siibak.** Effect of antibiotics on ribosome assembly is indirect. Tartu, 2011, 134 p.
196. **Tambet Tõnissoo.** Identification and molecular analysis of the role of guanine nucleotide exchange factor RIC-8 in mouse development and neural function. Tartu, 2011, 110 p.
197. **Helin Räägel.** Multiple faces of cell-penetrating peptides – their intracellular trafficking, stability and endosomal escape during protein transduction. Tartu, 2011, 161 p.
198. **Andres Jaanus.** Phytoplankton in Estonian coastal waters – variability, trends and response to environmental pressures. Tartu, 2011, 157 p.
199. **Tiit Nikopensius.** Genetic predisposition to nonsyndromic orofacial clefts. Tartu, 2011, 152 p.
200. **Signe Värv.** Studies on the mechanisms of RNA polymerase II-dependent transcription elongation. Tartu, 2011, 108 p.
201. **Kristjan Välk.** Gene expression profiling and genome-wide association studies of non-small cell lung cancer. Tartu, 2011, 98 p.
202. **Arno Põllumäe.** Spatio-temporal patterns of native and invasive zooplankton species under changing climate and eutrophication conditions. Tartu, 2011, 153 p.
203. **Egle Tammeleht.** Brown bear (*Ursus arctos*) population structure, demographic processes and variations in diet in northern Eurasia. Tartu, 2011, 143 p.
205. **Teele Jairus.** Species composition and host preference among ectomycorrhizal fungi in Australian and African ecosystems. Tartu, 2011, 106 p.
206. **Kessy Abarenkov.** PlutoF – cloud database and computing services supporting biological research. Tartu, 2011, 125 p.

207. **Marina Grigorova.** Fine-scale genetic variation of follicle-stimulating hormone beta-subunit coding gene (*FSHB*) and its association with reproductive health. Tartu, 2011, 184 p.
208. **Anu Tiitsaar.** The effects of predation risk and habitat history on butterfly communities. Tartu, 2011, 97 p.
209. **Elin Sild.** Oxidative defences in immunoecological context: validation and application of assays for nitric oxide production and oxidative burst in a wild passerine. Tartu, 2011, 105 p.
210. **Irja Saar**. The taxonomy and phylogeny of the genera *Cystoderma* and *Cystodermella* (Agaricales, Fungi). Tartu, 2012, 167 p.
211. **Pauli Saag.** Natural variation in plumage bacterial assemblages in two wild breeding passerines. Tartu, 2012, 113 p.
212. **Aleksei Lulla.** Alphaviral nonstructural protease and its polyprotein substrate: arrangements for the perfect marriage. Tartu, 2012, 143 p.
213. **Mari Järve.** Different genetic perspectives on human history in Europe and the Caucasus: the stories told by uniparental and autosomal markers. Tartu, 2012, 119 p.
214. **Ott Scheler**. The application of tmRNA as a marker molecule in bacterial diagnostics using microarray and biosensor technology. Tartu, 2012, 93 p.
215. **Anna Balikova**. Studies on the functions of tumor-associated mucin-like leukosialin (CD43) in human cancer cells. Tartu, 2012, 129 p.
216. **Triinu Kõressaar.** Improvement of PCR primer design for detection of prokaryotic species. Tartu, 2012, 83 p.
217. **Tuul Sepp.** Hematological health state indices of greenfinches: sources of individual variation and responses to immune system manipulation. Tartu, 2012, 117 p.
218. **Rya Ero.** Modifier view of the bacterial ribosome. Tartu, 2012, 146 p.
219. **Mohammad Bahram.** Biogeography of ectomycorrhizal fungi across different spatial scales. Tartu, 2012, 165 p.
220. **Annely Lorents.** Overcoming the plasma membrane barrier: uptake of amphipathic cell-penetrating peptides induces influx of calcium ions and downstream responses. Tartu, 2012, 113 p.
221. **Katrin Männik.** Exploring the genomics of cognitive impairment: whole-genome SNP genotyping experience in Estonian patients and general population. Tartu, 2012, 171 p.
222. **Marko Prous.** Taxonomy and phylogeny of the sawfly genus *Empria* (Hymenoptera, Tenthredinidae). Tartu, 2012, 192 p.
223. **Triinu Visnapuu.** Levansucrases encoded in the genome of *Pseudomonas syringae* pv. tomato DC3000: heterologous expression, biochemical characterization, mutational analysis and spectrum of polymerization products. Tartu, 2012, 160 p.
224. **Nele Tamberg.** Studies on Semliki Forest virus replication and pathogenesis. Tartu, 2012, 109 p.
225. **Tõnu Esko.** Novel applications of SNP array data in the analysis of the genetic structure of Europeans and in genetic association studies. Tartu, 2012, 149 p.

226. **Timo Arula.** Ecology of early life-history stages of herring *Clupea harengus membras* in the northeastern Baltic Sea. Tartu, 2012, 143 p.
227. **Inga Hiiesalu.** Belowground plant diversity and coexistence patterns in grassland ecosystems. Tartu, 2012, 130 p.
228. **Kadri Koorem.** The influence of abiotic and biotic factors on small-scale plant community patterns and regeneration in boreonemoral forest. Tartu, 2012, 114 p.
229. **Liis Andresen.** Regulation of virulence in plant-pathogenic pectobacteria. Tartu, 2012, 122 p.
230. **Kaupo Kohv.** The direct and indirect effects of management on boreal forest structure and field layer vegetation. Tartu, 2012, 124 p.
231. **Mart Jüssi.** Living on an edge: landlocked seals in changing climate. Tartu, 2012, 114 p.
232. **Riina Klais.** Phytoplankton trends in the Baltic Sea. Tartu, 2012, 136 p.
233. **Rauno Veeroja.** Effects of winter weather, population density and timing of reproduction on life-history traits and population dynamics of moose (*Alces alces*) in Estonia. Tartu, 2012, 92 p.
234. **Marju Keis.** Brown bear (*Ursus arctos*) phylogeography in northern Eurasia. Tartu, 2013, 142 p.
235. **Sergei Põlme.** Biogeography and ecology of *alnus*- associated ectomycorrhizal fungi – from regional to global scale. Tartu, 2013, 90 p.
236. **Liis Uusküla.**. Tartu, 2013, 173 p.
237. **Marko Lõoke.** Studies on DNA replication initiation in *Saccharomyces cerevisiae.* Tartu, 2013, 112 p.
238. **Anne Aan.** Light- and nitrogen-use and biomass allocation along productivity gradients in multilayer plant communities. Tartu, 2013, 127 p.
239. **Heidi Tamm.** Comprehending phylogenetic diversity – case studies in three groups of ascomycetes. Tartu, 2013, 136 p.
240. **Liina Kangur.** High-Pressure Spectroscopy Study of Chromophore-Binding Hydrogen Bonds in Light-Harvesting Complexes of Photosynthetic Bacteria. Tartu, 2013, 150 p.
241. **Margus Leppik.** Substrate specificity of the multisite specific pseudouridine synthase RluD. Tartu, 2013, 111 p.
242. **Lauris Kaplinski.** The application of oligonucleotide hybridization model for PCR and microarray optimization. Tartu, 2013, 103 p.
243. **Merli Pärnoja.** Patterns of macrophyte distribution and productivity in coastal ecosystems: effect of abiotic and biotic forcing. Tartu, 2013, 155 p.
244. **Tõnu Margus.** Distribution and phylogeny of the bacterial translational GTPases and the Mqsr/YgiT regulatory system. Tartu, 2013, 126 p.
245. **Pille Mänd**. Light use capacity and carbon and nitrogen budget of plants: remote assessment and physiological determinants. Tartu, 2013, 128 p.
246. **Mario Plaas.** Animal model of Wolfram Syndrome in mice: behavioural, biochemical and psychopharmacological characterization. Tartu, 2013, 144 p.