Migrations and Transfers in Prehistory: Asian and Oceanic Ethnolinguistic Phylogeography

University of Bern, 27 to 31 July 2014

Book of Abstracts
Contents

Evidence of Long-distance Dispersal events during the settlement of the world by modern humans
Isabel Alves¹,²,³, Miguel Arenas¹,²,⁴, Mathias Currat⁶, Anna Sramkova Hanulova¹,², Vitor Sousa¹,², Nicolas Ray⁵, Laurent Excoffier¹,² ................................................................. 1

Pakistan: An ethnolinguistic and genetic perspective
Qasim Ayub¹ and Chris Tyler-Smith² ................................................................. 2

Yenisseian homeland and migration
Václav Blažek (Masaryk University, Brno, Czech Republic) .................................................. 3

Introduction, Adoption and Rejection: Notes from Southeast Asian Archaeobotany
Cristina Castillo (University College London) ...................................................................... 4

The archaeogenetics of the Gangetic plain
Gyaneshwer Chaubey (Estonian Biocentre, Tartu) .............................................................. 5

Multiple dispersal model of modern human settlement in East Asia supported by statistical analyses and computer simulations of HLA genes
Di Da¹, Mathias Currat² and Alicia Sanchez-Mazas³ ................................................................ 6

Asian ethnolinguistic phylogeography from a linguistic perspective
George van Driem (University of Berne) .............................................................................. 7

The historical demography of Oceania as discerned from genetic data of modern population structure
Ana T. Duggan and Mark Stoneking (Max Planck Institute for Evolutionary Anthropology, Leipzig) .................................................................................................................. 8

Correlating Palaeo-Siberian languages and populations: recent advances in the Uralo-Siberian hypothesis
Michael Fortescue (University of Copenhagen and St Hugh’s College, Oxford) ......................... 9

Dating proto-languages through archaeological material correlates and imputing the immaterial for archaeological cultures: examples from Dravidian South India
Dorian Fuller (Institute of Archaeology, University College London) ......................................... 10

Correlations between paternal lineages and language families: On the cases of N-M231 with Sino-Uralic and O-M175 with Austic
Gao Jingyi (Institute of the Estonian Language, Tallinn) .......................................................... 11

Ancient Population Migrations in Northeast India: A closer look at the Ethnolinguistic Prehistory
Manjil Hazarika (Institute of Palaeoenvironment and Heritage Conservation, Mekelle University, Ethiopia, and Institut für Sprachwissenschaft, Universität Bern) ........................................................................... 12

Maritime contacts in the ancient Bay of Bengal: loanwords and sociolinguistic hierarchies
Tom Hoogervorst (Royal Dutch Institute of Southeast Asian and Caribbean Studies, Leiden) ...... 13
Associating Morphosyntactic Changes in Austronesian Languages with Human-migration Routes
Kikusawa Ritsuko (National Museum of Ethnology at Osaka and the Graduate University for Advanced Studies, Japan)......................................................................................... 14

Mapping Chromosome diversity in Eurasia
Toomas Kivisild (University of Cambridge) ........................................................................... 15

Language and Genes of the Greater Himalayan Region: mapping and correlating linguistic, genetic, and geographic barriers
P. de Knijff, G. Barbujani, D.R. Carvalho-Silva, G. van Driem, M.A. Jobling, Karma Tshering of Gaselö, T. Kraayenbrink, E.J. Parkin, Nirmal Man Tuladhar, C. Tyler-Smith ............................................. 16

Intricate history of physical groups, ethnic clusters, and linguistic families in East Asia viewed from Y Chromosomes
Li Hui (Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai) .................................................................... 17

Phylogeography, ethnobotany, and linguistics: issues arising from research on the natural and cultural history of taro, Colocasia esculenta (L) Schott
Peter J. Matthews, Peter L. Lockhart and Ibrar Ahmed ............................................................. 18

The population structure of Helicobacter pylori across Asia offers insights into prehistoric human migrations
Yoshan Moodley (Konrad Lorenz Institute, University of Veterinary Medicine Vienna) ............... 19

A preliminary comparison of linguistic and genetic diversity in Western Asia
Luca Pagani, Eugenio Bortolini, Giuseppe Longobardi, Giovanni Romeo, Donata Luiselli, Davide Pettener ........................................................................................................................................ 20

Identifying Prehistoric Population Trajectories: Who Influenced Iraya?
Lawrence A. Reid (University of Hawai‘i) ...................................................................................... 21

Proto-Trans-Eurasian: where and when?
Martine Robbeets (Johannes Gutenberg Universität Mainz) ........................................................... 22

Genomic DNA analyses of modern and ancient people on Japanese Archipelago
Saitou Naruya (Department of Population Genetics, National Institute of Genetics, Mishima, Japan & Department of Genetics, Graduate University for Advanced Studies, Department of Biological Sciences, University of Tōkyō) ........................................................................... 24

The linguistic area of Wallacea
Antoinette Schapper (Koninklijk Instituut voor Taal-, Land- en Volkenkunde & Universität zu Köln) ................................................................................................................................. 25
Evidence of Long-distance Dispersal events during the settlement of the world by modern humans
Isabel Alves\textsuperscript{1,2,3}, Miguel Arenas\textsuperscript{1,2,4}, Mathias Currat\textsuperscript{6}, Anna Sramkova Hanulova\textsuperscript{1,2}, Vitor Sousa\textsuperscript{1,2}, Nicolas Ray\textsuperscript{5}, Laurent Excoffier\textsuperscript{1,2}

\textsuperscript{1}CMPG, Institute of Ecology and Evolution, Baltzerstrasse 6, 3012 Berne, Switzerland
\textsuperscript{2}Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland
\textsuperscript{3}Population and Conservation Genetics Group, Instituto Gulbenkian de Ciência, Oeiras, Portugal
\textsuperscript{4}Centre for Molecular Biology “Severo Ochoa”, Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain.
\textsuperscript{5}Anthropology, Genetics and Peopling History Lab, Anthropology Unit – Department of Genetics & Evolution, University of Geneva, Geneva, Switzerland.
\textsuperscript{6}EnviroSPACE Lab., Institute for Environmental Sciences, University of Geneva, Battelle – Building D, 7 route de Drize, 1227 Carouge, Switzerland.

Genetic data have proved crucial to reconstruct the evolutionary history of human populations. Nevertheless, important demographic and historical aspects of this history have been neglected. For instance, the Last Glacial Maximum (LGM) has certainly promoted range contractions and range shifts of several species including humans and long-distance dispersal (LDD) plays an important role in many species such as plants, birds and fish and could have also been important in humans.

Therefore, we aim at testing whether the LGM and LDD have played a role in shaping currents patterns of genetic diversity of our species by statistically evaluating the relative probability of alternative models of the colonization of the old world by humans using extensive spatially explicit simulations embedded into an approximate Bayesian computation framework. We thus generated genetic data under four different models: I) a simple range expansion that has been used to describe the demographic history of humans; II) a range expansion followed by a range contraction mimicking the effect of the LGM; III) a range expansion with LDD; and finally IV) a model of expansion integrating both a LGM contraction and LDD events. We then compared the patterns of diversity and differentiation with a corresponding genotype dataset consisting of 50 microsatellites genotyped in 22 worldwide human populations. By applying a Bayesian model choice, we find that models with LDD are vastly better supported by the data, and that the best supported model includes both a LGM contraction and LDD events. This latter model can also correctly reproduce all observed patterns of diversity. We then discuss the validity and importance of these results for future modelling of human populations.
Pakistan lies in a region of south Asia that is now populated by almost 2 billion people representing many diverse ethnicities and linguistic groups. Despite lack of fossil evidence, archaeological evidence indicates that modern humans arrived here soon after they left Africa ~50 – 70,000 years before present (YBP), most likely by a southern coastal route. By 7,000 YBP they had started establishing cities that eventually expanded to represent the Indus Valley Civilization and Harappan culture that rivalled Mesopotamia. Since then several subsequent human migrations and numerous invasions, mostly from the west, have helped shape the pattern of genetic diversity currently observed in these populations. Studies utilizing markers on the human Y chromosome or mitochondrial DNA in selected populations provided initial glimpses into their male and female ancestry, respectively. These markers exhibited a mixture of Western Eurasian, South and East Asian specific lineages, some of which were unequivocally associated with specific past invasions. These and subsequent studies confirmed that in Pakistan genetic relationships are generally accurately predicted by geographic proximity rather than linguistic origin. The Dravidian-speaking Brahui population are a prime example of this. They currently reside in south western Pakistan, surrounded by Indo Europeans speakers with whom they share a common genetic origin. The availability of whole genome sequences from populations in this region will further assist in understanding fine scale population stratification. We aim to use low coverage sequenced South Asian samples from the 1,000 Genomes Project, and high coverage sequence of a Kalash individual to improve our understanding of the arrival and spread of Indo-European speakers in this region and unravel the enigma of the Brahuis.
Yenisseian homeland and migration

Václav Blažek (Masaryk University, Brno, Czech Republic)

The historical extension of speakers of the Yenisseian languages described in the last three centuries is limited to the basin of the Yenissei from the Kureika river in the north to the Kan river in the south. Taking into account the witness of toponyms, especially hydronyms, the southern border of the Yenisseian dialect continuum should be shifted to the basins of the rivers Abakan in Khakasia and Bolšoi Yenissei in North Tuva (Dul’zon 1959, 97; 1963, 289-95). The western border has usually been defined by the Irtyš/Ertis river. This Turkicized hydronym, known also in the Mongolian mediation as Ircis, consists of two components; the second may easily be identified with Common Yenisseian *ses "river, brook" (reconstruction - see Starostin 1995, 272; identification - see already Ramstedt 1907, 4).

Typical Yenisseian hydronyms also appear between the Demyanka and Tara rivers, both right tributaries of the Irtyš (Dul’zon 1959, 97). But there are traces of Yeniseian hydronyms even to the west of the Irtyš, namely the Selety river emptying into the lake Selety-Tengiz from Northern Kazakhstan, with the first component etymologizable as "stag", cf. South Ket šel’, Kott šele, Pumpokol salat ‘rangifer’ (see Blažek 1995, 9-10; proto-Yeniseian *sēr1e by Starostin 1995, 272).

These hydronymical traces indicate that the Yenisseian homeland was situated in the deeper past much more to the southwest, probably in the territory of Kazakhstan. The trajectory of the migration of early Yenisseians went probably along the 52 parallel of latitude, from North-Central Kazakhstan to Irtyš and further to the upper streams of the Ob and Yenissei rivers. The last river became a new axis of their northward migration.

References


Introduction, Adoption and Rejection: Notes from Southeast Asian Archaeobotany
Cristina Castillo (University College London)

Archaeobotany is revealing information on the movements of crops into and out of Southeast Asia. This paper will present some of the findings from current archaeobotanical research in mainland and island Southeast Asia with a focus on Peninsular Thailand. The results from the sites Khao Sam Kaeo and Phu Khao Tong situated in the Southern Peninsula of Thailand indicate the existence of economically important crops originating from Africa and South Asia from as early as the Metal Age. The importance of these crops in relation to their trajectory of likely dispersal from centres of origin, possible reasons for their movements and issues on food preference will be discussed in this paper. The existence of non-indigenous plants in the Southeast Asian archaeological record provides information on movements of people as well as trade during the early prehistoric period.
The archaeogenetics of the Gangetic plain
Gyaneshwer Chaubey (Estonian Biocentre, Tartu)

Extending between the Indus river valley in West and Brahmaputra river valley in East, the Gangetic plain is considered as the most densely populated region of the South Asia. Topographically it is divided into Upper, Middle and Lower Gangetic plains. It is a home of more than 100 ethnic groups who mainly speak various dialects of Indo-Aryan languages. Along with them, there are few minor groups affiliated to Austroasiatic, Dravidian and Tibeto-Burman language families. Moreover, the highly ripped sociocultural diversity transformed it into a multifaceted region. Until now the archaeogenetic studies of Gangetic plain is immature. Therefore, with the help of sociocultural data collected from Brahmins and Kshatriya (the upper caste), we evaluated the concordance and discordance of gene-culture coevolution by using all three genetic systems (mtDNA, Y chromosome and autosomes). We have also studied the genetic components of the Tharu, a malaria resistant population from Terai region.
Multiple dispersal model of modern human settlement in East Asia supported by statistical analyses and computer simulations of HLA genes

Di Da\textsuperscript{1}, Mathias Currat\textsuperscript{2} and Alicia Sanchez-Mazas\textsuperscript{3}

\textsuperscript{1} Laboratory of Anthropology, Genetics and Peopling history (AGP lab)

\textsuperscript{2} Department of Genetics and Evolution - Anthropology Unit, University of Geneva

\textsuperscript{3} Institute of Genetics and Genomics in Geneva (IGE3)

Genetic differences between northern (NEA) and southern (SEA) East Asian populations have been observed in numerous studies. In acceptance of their common origin in East Africa or West Asia, most of the debates are around the modern human migration route(s) into East Asia. In this study, we analyzed a large dataset of Human Leukocyte Antigen (HLA) system to reveal the East Asian genetic patterns, based on which a spatially explicit computer simulation approach was applied to investigate the main models for the human settlement of the region. We observed a genetic differentiation between NEA and SEA populations following a continuous pattern from north to south, and we remarked a significant and continuous decrease of HLA diversity by the same direction. This continuity is shaped by clinal distributions of many HLA lineages and alleles with increasing or decreasing frequencies along the latitude. These results bring evidence in favor of an “Overlapping model” we proposed, whereby modern humans migrated eastward from western Eurasia via two independent routes along either side of the Himalayas, and later overlapped in East Asia across open land areas. Our simulation results strongly support this model, emphasizing long-lasting gene flow between populations originating from the two routes probably since a remote prehistoric period, and recent emergence of a genetic boundary around the Qinling-Huaihe line or Yangtze River possibly linked to political and/or linguistic subdivisions. Moreover, our study suggests distinct demographic histories of NEA and SEA populations, characterized by severe isolation in the south. Finally, though a significant level of balancing selection acting on the three HLA loci was detected, its effect on the genetic patterns appears to be minor compared to those of past demographic events.
The Himalayan range runs the length of over 3600 kilometres from the Hazārahjāt Highlands in the west to the Liángshān in the east. The Himalayas form no natural watershed and many of the rivers are older than the mountains themselves. The Kali Gandaki River bisects the range into two halves of roughly equal length. The Eastern Himalaya runs from the Dhaulagiri massif across the Himalayas and sub-Himalayas, the Meghalaya, the lower Brahmaputra basin and associated hill tracts, the eastern Tibetan plateau and the Indo-Burmese borderlands into the provinces of Yúnnán and Síchuān.

The Eastern Himalaya is a region of pivotal importance in population prehistory. This Himalayan region served as a staging area and as a principal thoroughfare in the populating of much of Asia following the emergence of anatomically modern humans in Africa. New insights from linguistics, genetics and archaeobotany enable us tentatively to reconstruct the founding dispersals of a number of major language families in Asia and Oceania. The Eastern Himalaya appears to have served as a cradle of ethnogenesis at different time depths in the past.
The historical demography of Oceania as discerned from genetic data of modern population structure

Ana T. Duggan and Mark Stoneking (Max Planck Institute for Evolutionary Anthropology, Leipzig)

Within the past five years the rapid decrease in the cost of sequencing has allowed genetic investigators to study population genetics at a wider and deeper level than was previously imaginable. However, full genome sequencing on a population-level scale is still prohibitively expensive. By sequencing uniparental markers, meaning those passed exclusively from either mother or father to child, such as mitochondrial DNA and the Y-chromosome, we can distinguish differences in movements of women and men between populations and infer maternal and paternal historical relationships. Furthermore, by studying single nucleotide polymorphisms (SNPs), single mutations within the genome, irrespective of the uniparental markers we can determine whether individuals have genetic heritage from two or more source populations and the relative proportion of ancestry from these groups which can then be used to conclude at what time in history these two groups met and over what span of time did they contribute to the new gene pool.

Oceania is known to have experienced at least two major migrations of modern humans, the first settlers gave rise to the populations we now classify as Papuan and presumably spent tens of thousands of years in relative isolation before the Austronesian migration swept through contributing an abundance of new languages and cultural complexes as well as significantly altering the gene pool. By studying the genetic makeup of present-day Oceanians we can identify the past relationships between these two groups of people. The genetic contribution of Austronesian women appears to be particularly great; however we also find evidence of deep genetic lineages in the region indicating a long settlement history with incomplete replacement and the composition of Polynesian populations suggests a greater Papuan male contribution than might be expected.

These three pieces of genetic evidence, exclusively maternal, exclusively paternal and those portions which are equally contributed by both parents, allows us to reconstruct the genetic composition of Oceanian populations both present and past. With these reconstructions we can then deduce in fine detail the history of populations, their movements within Oceania, their interactions with each other and the approximate timing of these events.
Correlating Palaeo-Siberian languages and populations: recent advances in the Uralo-Siberian hypothesis
Michael Fortescue (University of Copenhagen and St Hugh’s College, Oxford)

Since the publication of Fortescue (1998), in which I presented a synoptic model for the origin and present distribution of languages across the Eurasian Arctic and into North America – the so-called Uralo-Siberian hypothesis – a number of new facts and theories within linguistics, archaeology and population genetics have come to light that necessitate adjustments to that model and to the overall picture of prehistoric population movements among ‘Palaeo-Siberians’. These newer sources include Irina Nikolaeva’s historical dictionary of Yukaghir (2006), my own comparative dictionary of Chukotko-Kamchatkan (2005) and subsequent focus on the relationship between Chukotoko-Kamchatkan and Nivkh (Fortescue 2011), also the work of Edward Vajda and colleagues indicative of a genetic link between Yeniseian and Na-Dene languages (Kari & Potter 2010). Advances in the rapidly developing picture of mtDNA-based genetic relationships between Arctic populations can be found in Volodko et al. (2008), for example.

My paper will concentrate on two sub-groups: Eskaleut-Yukaghir (considered independently of the broader Uralic link), and Nivkh-Chukotko-Kamchatkan, which I shall attempt to fit into a broader picture of population movements and expansions in the region. The alignment of genetic populations and linguistic families still presents considerable difficulties, complicated by potential population mixing and language shift amongst small mobile groups of hunter-gatherers, such as typified late Pleistocene and Holocene times in Siberia and Beringia. I shall nevertheless tentatively suggest a new scenario for the correlation of population genetics and language in the circum-Beringian world based on overlapping waves of migration towards and into the New World.


Historical linguistics has well-established comparative methods for inferring the phylogenetic relationships of languages and thus inferred shared proto-languages, with at least partly known vocabularies and thus concepts about the world. However, these proto-languages float in both space and time, two areas at which archaeology excels. All archaeological is fixed in space, at sites and regional settlement patterns, and in time, through stratigraphy and radiometric dating. By comparing these two sources of evidence there is high potential for delimiting chronological and geographical space of particular proto-languages, which will be illustrated from examples in the Dravidian language family (Proto-South Dravidian, Proto-South-Central Dravidian, Proto-Dravidian) by considering trees, crops, animals, and material technologies such as ceramics, metallurgy and textile production. This also allows us to place introduced crops and long-distance cultural contacts into a clear chronological framework. In some cases this is evident in the archaeology and in other cases it implied by the linguistics. In addition, what this reveals is that archaeology allows us to posit additional sub-phases between Proto-South-Central and Proto-South Dravidian than are evident from linguistic evidence (perhaps they would be if we had access to additional presumably extinct language branches). However, having fixed these proto-languages (or language clades) in time and space, this provides the opportunity to use less material concept restricted from historical linguistics (such as kings, kinship and marital systems, potential mytho-religious concepts) to fill out our understanding of past cultures beyond what is strictly observable from archaeology. The potential to look at cooking styles both through archaeological and historical linguistics will also be considered.
Correlations between paternal lineages and language families: On the cases of N-M231 with Sino-Uralic and O-M175 with Austric

Gao Jingyi (Institute of the Estonian Language, Tallinn)

Several geneticists have correlated their paternal lineages (human Y-chromosome haplogroups) with the existing theories of language families without any proposition in linguistics. The present author, as a linguist focusing on etymology, correlates the paternal lineages with ongoing researches of etymological relations, and finds more suitable correlations, which support the Father Tongue hypothesis elaborated by the historical linguist van Driem (2010).

The paternal lineage N-M231 has been correlated with the Uralic (=Finno-Ugric) language family (Zerial et al. 1997, Karefet et al. 2002). However, it has been proven that this paternal lineage arose in the Far East (Rootsi et al. 2007, Shi et al. 2013), or directly speaking, in China. Should the correlation between N-M231 and Uralic be abandoned? Most people think that there is no way to include Chinese to this correlation.

The author had independently identified etymologies shared by Sinitic and Uralic languages before knowing anything about population genetics. The etymological relation between Sinitic and Uralic languages has been recently well demonstrated (Gao 2014). In fact, the author has suggested the correlation between N-M231 and Sino-Uralic (=Sino-Finnic) in his first monograph in Chinese (Gāo 2008: 104–108).

Not O-M175 but N-M231 is the foundation of Han Chinese, because:

(1) N-M231 is more prevalent in western and central China, especially in the area of Yangshao culture, where was the cradle of the Chinese civilization. O-M175 is evenly prevalent in all Chinese areas, including those which were not historically Chinese.

(2) N-M231 is also prevalent in several Tibeto-Burman populations and most Uralic populations, whose linguistic relationship with Sinitic is becoming more and more evidential. O-M175 is much more prevalent in Hmong-Mien, Tai-Kadai, and Austronesian populations, whose linguistic relationship with Sinitic is becoming more and more controversial.

(3) Hlai (Tai-Kadai), She (Hmong-Mien) and Formosan (Austronesian) populations are up to 100% O-M175. Nenets (Uralic) and Nganasan (Uralic) populations are up to 97.3% N-M231. It is reasonable to suggest that proto-Austric was 100% O-M175, while proto-Sino-Uralic was 100% N-M231. The high appearance of O-M175 in Han Chinese populations is due to language shifts from Austric to Chinese.
Ancient Population Migrations in Northeast India: A closer look at the Ethnolinguistic Prehistory
Manjil Hazarika (Institute of Palaeoenvironment and Heritage Conservation, Mekelle University, Ethiopia, and Institut für Sprachwissenschaft, Universität Bern)

Four major linguistic families are found in Northeast India, i.e. Austroasiatic, Tibeto-Burman, Indo-European and Kradai. Previous studies on the reconstruction of the sequence of arrival of these people of different linguistic stocks have been based on simple philological and ethnographical considerations. The unfolding story revealed by population genetics and the linguistic palaeontology of the different language families turns out to be more complex than previously thought. The present paper aims at a fresh investigation of ethnolinguistic prehistory of Northeast India based on archaeological and relevant multidisciplinary data. The evidence presented in this paper is gathered from archaeobotanical, ecological, ethnobiological, ethnographical, and genetic sciences to inspire an interpretation of the available archaeological data for examining linguistic hypotheses of ancient population migrations and dispersals in this region.
The study of loanwords is a key tool to reconstruct cultural contact and understand its multifaceted nature. It is also fraught with difficulties, as we are often dealing with historical languages only partly known to us. This paper examines processes of interethnic contact and lexical borrowing across the ancient Bay of Bengal. The archaeological record presents a picture of intensified contact between India’s east coast and the Malay World from the mid-first millennium BCE to the late first millennium CE, after which Islamic and – later – colonial networks start to replace the earlier dynamics that form the focus of this paper. It is my aim here to cast a new light on the speech communities from either side of the Bay of Bengal that were involved in the maritime exchange of commodities, words and ideas. This involves sharpening our picture of the ethno-linguistic macrohistory of this phylogenetically diverse area, in particular of the historico-phonological development of the Dravidian, Indo-Aryan and West-Malayo-Polynesian languages under research. Complicated systems of social hierarchy add another dimension to this story. It is demonstrated that esteemed literary languages such as Sanskrit display various ways of masquerading foreign loans, which were seen as corruptions (apabhraṃśa) to the language’s perceived purity. Conversely, the success – if not raison d’être – of trade-oriented vernaculars centred on their susceptibility to external influence. Speakers of Malayic languages, hence, display a tendency to substitute inherited vocabulary by loanwords from Sanskrit, Arabic or other “high” languages, reflecting a deliberate attempt to manoeuvre themselves into the orbit of more cosmopolitan nodes of interaction.
Associating Morphosyntactic Changes in Austronesian Languages with Human-migration Routes
Kikusawa Ritsuko (National Museum of Ethnology at Ōsaka and the Graduate University for Advanced Studies, Japan)

In this presentation, I will attempt to relate morphosyntactic changes proposed for Austronesian languages to the currently accepted migration hypothesis of the Pacific and discuss what kind of additional hypotheses could be made.

Linguistic subgrouping hypotheses have played an important role in identifying peoples’ migration routes of early speakers of Austronesian languages. The major phenomena looked at in such work are sound correspondences and lexical innovations where the Comparative Method is applied. However, other linguistic aspects, such as morphosyntax and semantics, also undergo change and their developmental paths should also be traceable. Thus a question arises as to the kind of contributions that are potentially expected in understanding peoples migration in prehistory when examining such changes.

I will discuss how the development of basic sentence structures and actancy systems proposed in Kikusawa 2003 as a result of macro-comparison of Austronesian languages relate to the existing migration theory. More specifically, I will show that, assuming that the given hypothesis is correct, morphosyntactic remnants appear to support rapid spread of Austronesians and possibly longer periods of incubating areal features in each macro-region.
Human genetic diversity in non-African populations derives predominantly from a recent dispersal from Africa. Previous Y chromosome studies have shown that haplogroups common throughout non-African populations all coalesce to a small number of shared ancestral lineages, the branching order of which is only partly understood. We have sequenced at high coverage the complete Y chromosomes of more than 200 individuals sampled widely across the world. We determined the regions of highest mapping quality and applied phylogenetic methods to determine the order and temporal dynamics of branching events in non-African Y chromosome haplogroups. Compared to the length of the branches that separate African and non-African diversity the internal branches distinguishing continental and sub-continental differences outside Africa are short and consistent with the model of a rapid colonization of Eurasia and Oceania. Bayesian skyline plots suggest that the signal of the initial phase of recovery and growth of Y chromosome diversity after the Out of Africa bottleneck has been best preserved in South and Southeast Asian populations. In contrast, extant Y chromosome diversity in Europe and Middle East coalesces to a small number of mid-Holocene founders majority of which share common ancestors with East Asian lineages at significantly younger dates than the initial Out of Africa radiation.


Language and Genes of the Greater Himalayan Region: mapping and correlating linguistic, genetic, and geographic barriers
P. de Knijff¹, G. Barbujani², D.R. Carvalho-Silva³, G. van Driem⁷, M.A. Jobling⁴, Karma Tshering of Gaselô⁵, T. Kraayenbrink¹, E.J. Parkin⁴, Nirmal Man Tuladhar⁶, C. Tyler-Smith⁵

¹ Leiden University
² Università di Ferrara
³ Wellcome Trust Sanger Institute, Hinxton, UK
⁴ Leicester University
⁵ Australian National University Canberra
⁶ Centre for Nepal and Asian Studies, Tribhuvan University
⁷ Institut de Linguistique, Université de Berne

There are a number of competing theories about the population history of the Himalayas. These theories are largely based on linguistic and/or archaeological findings, sometimes supported by the results of small-scale genetic studies. A large-scale, ethnolinguistically informed, genetic study of the greater Himalayan region might provide a definite model for historical population events in this region, which is why the current study was initiated.

Bhutan and Nepal could have been corridors for human migration through the Himalayas in ancient times. Depending on the direction of migration these countries could be one of the oldest inhabited Himalayan regions, a theory that is supported by the multitude of language communities. A genetic study of the Bhutan and Nepal, therefore, may not only provide evidence for the uniqueness and antiquity of residence people in the Himalayan region, but may also clarify the discussion about the peopling of the Himalayas and Eastern Asia in general.

Here I will summarize the results of our study that was aimed at answering the following questions:

1. Is there a correlation between language and genes in the Himalayan region?
2. Can we determine the genetic relationships (ancient ancestors) of the Nepalese and Bhutanese and deduce possible migration routes?
3. Can we say something about relative ages of the various groups, comparing the few real "aboriginal" groups with the others?
Intricate history of physical groups, ethnic clusters, and linguistic families in East Asia viewed from Y Chromosomes

Li Hui (Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai)

East Asia harbors substantial genetic, physical, cultural, and linguistic diversity, but the detailed structures and interrelationships of those aspects remain enigmatic. This question has begun to be addressed by a rapid accumulation of molecular anthropological studies of the populations in and around East Asia, especially by the Y chromosome studies. The current Y chromosome evidences suggest multiple early migrations of modern humans from Africa via Southeast Asia to East Asia involving so called Australian, Negrito, and Mongolian styles of modern humans. After the initial settlements, the northward migrations during Paleolithic Age shaped the genetic structure in East Asia. Subsequently, recent admixtures between Central Asian immigrants and northern East Asians enlarged the genetic divergence between southern and northern East Asia populations. Ethnic clusters and Linguistic families formed along with the archaeological cultures during the Neolithic Age, e.g., Sinitic subfamily with Hongshan and Longshan Culture contexts, Tibeto-Burman with Yangshao Culture context, Hmong-Mien with Daxi Culture context, Tai-Kadai and Austronesian with Liangzhu Culture context. A southward migration of Austronesians also started from mainland southern East Asia, travelled through Malaysia and Indonesia, and eventually navigated eastwards to the Pacific Islands.
Phylogeography, ethnobotany, and linguistics: issues arising from research on the natural and cultural history of taro, *Colocasia esculenta* (L) Schott

Peter J. Matthews¹, Peter L. Lockhart² and Ibrar Ahmed²,³

¹ National Museum of Ethnology, Japan
² Massey University, New Zealand
³ Quaid-i-Azam University, Pakistan

Ever since the emergence of modern genetics in the early to mid-20th century, there have been scattered attempts to explore the genetic and geographical origins of cultivated taro. There have also been discussions of how local names for the plant in different languages might reflect the origins and movements of cultivated taro. Early cytological studies were consistent with botanical and linguistic evidence indicating that the species originated in Asia, but the question of how to recognise the natural range of taro has always created difficulty.

The ability of taro to self-propagate vegetatively, escape from cultivation, and become feral or naturalised, has always been obvious. Less obvious has been the fact that in tropical and subtropical regions, flowering and breeding of wild taro populations is common, whatever their natural or cultural origins may be. Recent ethnobotanical research suggests that the process of naturalisation has been widely encouraged by deliberate planting in suitable wild or modified (but not cultivated) habitats. Under these circumstances, how is it possible to recognise naturally distributed wildtype forms, and thus the natural geographical range within which first use and early or primary domestication might have taken place?

Our recent study of chloroplast DNA diversity in taro and closely related species provides a partial solution to the problem, and raises new questions regarding the origins of cultivated taro. Our new model for the origins of cultivated taro can be related to previous suggestions based on linguistic data. Together, the ethnobotanical and genetic studies also raise questions about what kinds of linguistic data are needed to investigate crop origins and dispersal history.

At a broad or high level of historical analysis, there is good reason to treat linguistic and biological data separately. At a local or ground level, human relationships with plants are mediated through local language, so the processes of biological and linguistic change cannot be independent. To understand these processes, and develop more realistic models of crop origins and dispersals, closer integration of linguistic, cultural, and biological research are needed. How people identify plants, use them, and move them about in landscapes is closely related to first-hand physical experience, and to behavioural or cultural traditions that are learned through seeing, doing, and language.
The population structure of Helicobacter pylori across Asia offers insights into prehistoric human migrations
Yoshan Moodley (Konrad Lorenz Institute, University of Veterinary Medicine, Vienna)

The ancient association of Helicobacter pylori with humans has led both species to very similar evolutionary histories. We investigated the processes involved in H. pylori’s colonization of the Asian continent by applying a spatially explicit migration model. Our analyses of 1164 strains from 69 locations and ethnicities throughout Asia showed the presence of three major populations. These included hpAsia2 and hpEastAsia, the latter of which is further subdivided into the subpopulations hspAmerind, hspEasia and hspMaori, but also the hitherto undiscovered hpSiberia. This population was at highest frequency in central Siberia and using a linkage model to determine ancestral nucleotide states, we identified it as a hybrid between the Central Asian hpAsia2 and the East Asian hspEasia. Furthermore, we detected the subpopulation hspAmerind, usually associated with indigenous Americans, at high frequency at its origin along the eastern Pacific seaboard, but also in central Siberia and as far west as the Yamal Peninsula. This suggests that humans not only carried hspAmerind from its origins eastwards into the Americas, but also travelled thousands of kilometers westwards to the very fringes of northern Europe. These analyses allow for a better understanding of the prehistoric human expansions that peopled the Asian continent.
A preliminary comparison of linguistic and genetic diversity in Western Asia
Luca Pagani1,2, Eugenio Bortolini2, Giuseppe Longobardi3, Giovanni Romeo2, Donata Luiselli2, Davide Pettener2

1 The Wellcome Trust Sanger Institute, Hinxton, Great Britain
2 Department of Biological, Geological and Environmental Science, University of Bologna, Italy
3 Department of Language and Linguistic Science, University of York, Great Britain

The area spanning from the Mediterranean Sea to the Indian Ocean hosted over the last six millennia many complex societies. Most of them left behind some of the earliest evidences of written language known to date. Across the whole area, remnants of a number of different writing systems have indeed been found over the years, documenting a long-standing occupation of the area by various human populations, certainly speaking different languages. Such ancient diversity is still detectable in present human populations. At least four linguistic families (Indo-European, Semitic, Altaic, and Dravidian) are now present in the area. Despite the recent diffusion of the Islamic package, linguistic diversity in the region is remarkably resilient and might still be a good predictor of diversity in cultural and genetic descent. Strikingly, people inhabiting the three biggest Islamic countries in the region (Turkey, Iran, Pakistan and Saudi Arabia) speak languages belonging to the Indo-Iranian (Farsi in Iran and Urdu in Pakistan), Altaic (Osmanli Turkish in Turkey), Dravidian (Brahui in Pakistan) and Semitic (Arabic in Saudi Arabia).

Such a fragmented scenario, encompassing a number of different cultural and linguistic packages in a relatively small geographic region, provides an ideal case to test the hypothesis of co-evolving cultural-linguistic and genetic patterns. This hypothesis, proven true in many cases including European (Longobardi et al 2013, submitted), Ethiopian (Pagani et al. 2012) and Himalayan (Chaubey et al. 2011) populations, states that linguistic differences could provide a genetic barrier between human populations.

In order to estimate linguistic distances between distantly related languages we relied on the Parametric Comparison Method (Longobardi and Guardiano 2009). By focussing on a set of discrete and universal syntactic traits, this approach has proven to be more robust to saturation than methods comparing lexical items. These cannot safely identify cognate sets across distinct families. Furthermore our method could potentially facilitate a better temporal resolution of linguistic events.

Linguistic distances hence estimated were compared with the genome-wide genetic data obtained from a set of populations belonging to the various linguistic groups present in the studied region. While a generally good correlation was observed, the Arabic/Farsi and Hebrew/Farsi pairs stood out as outliers. In particular, the linguistic distance between these populations exceeded the one predicted by their genetic distance, flagging the possibility of interesting mechanisms of cultural transmission/retention over time.
Identifying Prehistoric Population Trajectories: Who Influenced Iraya?
Lawrence A. Reid (University of Hawai‘i)

Iraya, an endangered language of the Philippines and spoken in the north of the island of Mindoro in the central Philippines is unique among Philippine languages in a wide variety of features. A number of these features appear to be the result of a massive influx of an outside population that intermarried with the original people and imposed their language on them, a scenario that is supported by some of the traditional stories told in the community. This paper provides an overview of the Iraya people and their current situation. It also outlines some of the unique features of their language and discusses the problem of identifying who the people were who must have influenced the language so drastically. This is a work in progress, and reports the tentative results of my recent research in Mindoro.
Proto-Trans-Eurasian: where and when?
Martine Robbeets (Johannes Gutenberg Universität Mainz)

The label “Trans-Eurasian” was coined by Johanson & Robbeets (2010: 1-2) to refer to a large group of geographically adjacent languages, traditionally known as “Altaic”, that include up to five different linguistic families: Japonic, Koreanic, Tungusic, Mongolic, and Turkic. The question of whether these families go back to a single common ancestor called “proto-Trans-Eurasian” is one of the most disputed issues in historical comparative linguistics. The controversy is not primarily fueled by a shortage of similarities, but by the difficulty of accounting for them: are all shared forms generated by vertical transmission (i.e. borrowing), or are some of them residues of horizontal transmission (i.e. inheritance)? In previous research, I have argued that a small core of similarities between these languages, including regular sound correspondences, basic vocabulary, verb roots, derivational and inflectional morphology can be attributed to a common ancestor (e.g. Robbeets 2005, 2007a/b, 2010, 2012). However, if this indeed is the case, where and when did the speakers of proto-Trans-Eurasian originally live? And, how, why and when did the family split up and did the daughter languages move to their present locations?

In this paper, I will first summarize the ethnohistorical evidence for situating the five original speech communities of Turkic, Mongolic, Tungusic, Koreanic and Japonic in a rather compact area comprising North Korea, Southern Manchuria and present-day Southeastern Mongolia between the first and second millennium BC. On the basis of extra-linguistic considerations, such as the Neolithic revolution in the region taking place before 5000 BC as opposed to the relatively late date of separation of the individual branches, Manchuria’s high language family density ratio and the early but single presence of the Hongshan cultural complex (ca. 5000-3000 BC) in the region, I will then propose the hypothesis that proto-Trans-Eurasian goes back to the times of the Neolithic revolution in Manchuria and that it can be associated with the Hongshan culture (Nelson 1995, Janhunen 1996).

Finally, I will propose a scenario for the expansion of the individual branches over Eurasia, paying special attention to a possible link between the migration of the Koreanic and Japonic languages and the spread of wet rice agriculture. Recent archaeobotanical studies such as Crawford & Lee (2003), Miyamoto (2009), Ahn (2010) have shown that the spread of wet rice agriculture to Korea and Japan begins in Shandong, crosses via the Liaodong peninsula, reaches the Korean peninsula around 1300 BC and the Japanese islands after 1000 BC. From a linguistic standpoint, this can be associated with the entry of the Japonic language family into the Korean peninsula, and subsequently with the advent of Japonic in the Japanese islands.

References


Japanese Archipelago stretches over 4000 km from north to south, and is homeland of three human populations; Ainu, Mainlander, and Ryukyuan. Origins of these people have been studied for long time. Standard theory based on craniofacial data is "dual structure model" proposed by Hanihara (1991). According to this model, first migrants to Japanese Archipelago came from somewhere in Southeast Asia in Upper Paleolithic age, who were ancestors of Jomon people. Second wave of migration took place later in Yayoi period (3000-1700 BP), and people came in this time from Northeast Asia. Indigenous Jomon people and new migrants in and after Yayoi period gradually mixed with each other. This model provides reasonable explanation for morphological similarity between Ainu people of Hokkaido, northernmost main island of Japanese Archipelago, and Ryukyuan (or Okinawan) people in Southwest Archipelago, despite of large geographical distance. Similarity of these peoples was already noticed more than one hundred years ago by von Baelz (1911) as the Ainu-Ryukyuan common origin theory. Japanese Archipelago Human Population Genetics Consortium (2012) determined genome-wide single-nucleotide polymorphisms (SNPs) for Ainu and Ryukyuan, and compared these with existing data sets. This was first report of these genome-wide SNP data. Major findings were: (1) Recent admixture with Mainlander was observed for more than one third of Ainu individuals from two different analyses; (2) Ainu population probably experienced admixture with another population, and combination of two types of admixtures is the unique characteristics of this population; (3) Ainu and Ryukyuan are tightly clustered with 100% bootstrap probability followed by Mainlander in phylogenetic tree of East Eurasian populations. These results clearly support dual structure model on Japanese Archipelago populations, though origins of Jomon and Yayoi people still remain to be solved. We examined nuclear genome sequences of Jomon people (Kanzawa Hideaki et al., unpublished), because those ancient DNAs are direct evidences of human history. Ainu people were found to be genetically closest to Jomon people among present-day populations, and Mainland Japanese also inherited significant proportion of Jomon genomes. Furthermore, Denisovan gene flow to Jomon ancestors was suspected through statistical analyses. We also started analysis of genomic diversity within Japanese Archipelago Mainlander, and it may be possible to divide second wave of migration into two.


The linguistic area of Wallacea
Antoinette Schapper (Koninklijk Instituut voor Taal-, Land- en Volkenkunde & Universität zu Köln)

“Wallacea” typically refers to a zoogeographical area constituting a transition zone between Sundaland (the Malay Peninsula, Sumatra, Borneo, Java, and Bali) and Sahul (Australia and New Guinea) that includes Sulawesi, Lombok, Sumbawa, Flores, Sumba, Timor, Halmahera, Buru, Seram, and many smaller islands of eastern Indonesia and independent Timor-Leste. In this paper, I synthesize numerous strands of recent linguistic work that supports seeing Wallacea also as a linguistic area characterized by multiple shared features in the numerous Austronesian and Papuan languages it is home to. I argue that the nature and dispersal of the features shared by Wallacean languages points to the existence of Pleistocene networks of seafaring agriculturists.