Final Report: Maanasa Raghavan, Gibbs Travelling Research Fellow (2017-2018)

Reconstructing the population history of India using ancient and modern genomics

Overall research goals:

I am a molecular biologist, specializing in the use of next-generation sequencing data to study human evolutionary histories. My research interests span questions and applications in multiple fields, including *population genetics, genomics*, *anthropology, archaeology* and, more recently, *medical genetics*. The big question driving my research is: 'How have demographic, cultural, and environmental factors contributed over time to shaping the demographic and disease profiles of present-day human populations?' In order to address specific research questions embedded within this overarching goal, I use a combined approach that brings together genome-scale data from present-day human populations as well as ancient specimen. Over the past two years, I have been working in South Asia and implementing novel genomics tools in India to be able to study the origins and evolution of humans in this region. South Asia is particularly interesting due to the vast genetic and cultural diversity and the ongoing unresolved debates surrounding the demographic histories of these diverse human populations. The Gibbs Travelling Fellowship formed a crucial pilot for me to start building networks with Indian archaeologists, anthropologists, geneticists, and indigenous communities, as well as generate preliminary data forming the basis for the reconstruction of the long-term population history of South Asia.

Research during my tenure as a Gibbs Fellow:

As a Gibbs Fellow, my primary aim was to lay the foundation for long-term research into reconstructing the complex human population history of South Asia, using both ancient and modern human genetics. Several studies have shown that genetic data from ancient individuals are indispensable when reconstructing past population histories. Importantly, this region's genetic past remains a subject of great debate due to numerous hypotheses surrounding population origins and migrations within and from outside, including the uptake of practices associated with the Neolithic Period i.e. agriculture and domestication (~9,000 years before present), spread of the Indo-European languages (~4-5,000 years before present), as well as historic conquests and population migration from various parts of Eurasia. Whether some of these early cultural shifts are associated with human migrations (rather than transfer of idea) and, if so, how have they impacted the present-day gene pool, is an open question. Moreover, recent studies have shown that the population structure, especially practices such as endogamy and processes like founder effect, also have implications for the health status of humans in this region, who carry a heavy burden of recessive disorders. Hence, disentangling the present-day population structure of South Asia and ascertaining how and when it was established is an important research avenue for different fields.

With the above aim in mind, I conducted fieldwork in India in 2017-2018 to bring together samples that would increase the genomic representation of this region in both the ancient and modern genetics literature. This work benefitted greatly from collaborations and advice from researchers and mentors spanning various fields such as archaeology, anthropology, genetics, and computational biology, as well as the valuable participation of local communities.

<u>Ancient samples</u> – In collaboration with archaeologists at Deccan College Pune and the University of Hyderabad, I was able to access numerous human teeth and petrous bone samples (both considered rich sources of ancient DNA) from ongoing and past archaeological excavations at sites across India. We focused on material dating between 2-4,000 years before present from sites such as Rakhigarhi (Harappa Civilization) and Binjor (Harappa Civilization) in northwestern India, and sites across Telangana and Andhra Pradesh in southern India. We also collected a few younger samples (last 1,000 years) from other

regions such as Kashmir and Northeast India to ensure continuity in the genetic dataset and representation across the country. Unfortunately, I did not have the opportunity to visit the ongoing excavation at Rakhigarhi due to scheduling issues (although I had the chance to visit this site previously in 2016 when I trained the students on proper sampling techniques for ancient DNA while minimizing contamination), but I was able to access the samples from recent digs as they were brought back to the respective institutions and catalogued (and, in some cases, CT-scanned prior to destructive analyses). Additionally, while visiting these institutions, I was invited to give seminars on the use of scientific tools in archaeological studies. One of these invitations at the Indian Museum, Kolkata, was coupled with a visit to the Head Office of the Anthropological Survey of India (AnSI), where I had the opportunity to meet with researchers to discuss the possibilities of studying their collections, constituting a vast diversity of human samples across time and space. This networking opportunity recently turned into a collaboration as my key collaborator in India, Dr. Niraj Rai (Birbal Sahni Institute of Palaesciences, BSIP, Lucknow) and I have been granted access to several samples in the AnSI collections for undertaking genetic analyses. Hence, what started off as a pilot involving up to 100 samples has resulted in a rapid expansion of the scope to include many more sub-projects and samples from various time periods and cultures. I would also like to add here that during my tenure as a Gibbs Fellow, I actively helped Dr. Rai in his efforts to establish a state-of-the-art ancient DNA laboratory in India, which materialized in spring 2018. This space at BSIP has quickly become an integral resource for both our groups as we have been undertaking sample processing and training here, which has resulted in several research opportunities for local staff and students that would not have been available had the samples been taken out of the country. Moreover, we have also partnered up with a commercial DNA sequencing facility in India, MedGenome Inc., where all the data generation takes place, once again ensuring that all key stages in the project are performed locally in India.

<u>Present-day samples</u> – Since the analysis and interpretation of the ancient dataset is contingent on having a representative present-day genomic database, I also undertook fieldwork to collect blood samples and associated phenotype and cultural information from five modern populations: Nairs (different regions of Kerala), Kodavas (Coorg, Karnataka), Kaplas (Coorg, Karnataka), Bunts (Mangalore, Karnataka), and the Christian and Hindu Konkani diaspora in Mangalore (Mangalore, Karnataka). The blood samples were collected in order to generate whole genome sequences, which provide an unbiased genetic resource (in comparison to the alternative SNP-chip datasets) to perform various population and medical genetics analyses. Ethical approval for this work was granted by the Human Biology Research Ethics Committee, University of Cambridge (Application number HBREC.2017.29). Upon advice from my Indian collaborators, I decided to focus on the southwestern region of India (Karnataka and Kerala) where they have already worked with communities before on other genetics-based projects and, hence, have established a long-term relationship. Prior to sampling, I also gave public talks and held discussion sessions at some of these locations to present our project aims and give the communities an insight into what genetics can tell us about population history and health status. This also gave the communities a chance to interact with me directly and only those who felt comfortable with the research aims were then invited to provide samples. There are several interesting populations residing in this region today, including urban dwellers and indigenous groups. Of our five populations, one is a transitioning hunter-gatherer group (Kaplas), while four are agriculturalists/urban dwellers – all with unique cultural and phenotypic traits. All sampled populations practice strict endogamy (i.e. marriage within the population/group). With the help of my collaborators, we collected blood as well as cultural and phenotypic information from all five populations. Examples of the latter information include: age, sex, village/town that parents and grandparents belonged to, language spoken, subsistence, diet, marriage patterns, eye and hair colour, height and weight, and melanin index (measured using a spectrometer). This information will ultimately be used to study the genetic versus environmental bases for inter-population phenotypic differences. We sampled approximately 100 individuals, which was significantly lower than what we had aimed for. However, we wanted to ensure that the study is conducted ethically and with utmost respect for these communities and, hence, we only sampled individuals who understood our aims

well and were still interested in participating. I will note that while undertaking this work we also started contacting other communities in this region as well as other parts of India, and will be undertaking outreach work before we extend invitations to more populations to collaborate with us in the near future.

Data generation – Since March 2018, I have been working in India with Dr. Rai to process the ancient and modern samples. We have successfully completed an initial round of sequencing of 80 out of the 100 modern samples as well as all 100 ancient samples, and have started the data analysis.

Ongoing and future work stemming from the Gibbs Fellowship:

Results from the ongoing data analysis are now being used to evaluate the feasibility of ancient genomics in India by assessing the long-term preservation of DNA in various parts of India. Samples that yield authentic ancient DNA will then be subject to further sequencing to get higher depth and coverage across the human genome. Ultimately, this data will be analyzed in the context of modern DNA sequences to understand the evolutionary trajectory of humans in South Asia and test for genetic versus cultural affinities with worldwide populations. The modern dataset by itself is expected to reveal key insights into population-specific genetic variants associated with diseases and adaptions. The ancient DNA dataset will also be mined to look for the presence of ancient pathogens at the time of the individual's death, which will provide us with the first insights into the disease load in ancient South Asia. With the resulting publication(s), we hope that we can garner more attention and interest from archaeologists and local communities to initiate collaborations and undertake fine-scale genetic analyses of past and modern populations in the region. Along with my existing academic collaborators, we will be taking our current results back to all interested communities to jointly discuss what this means for them and how insights into their disease load can eventually be further developed for use by medical practitioners in diagnosis and treatment.

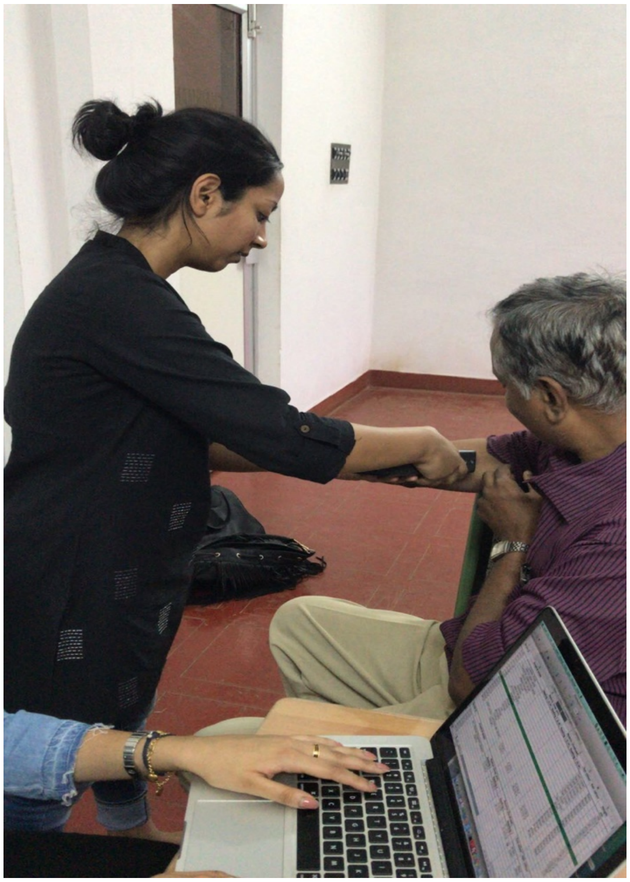
Acknowledgements

I am indebted to the Gibbs Fellowship Committee and Newnham College for supporting my research aims and equipping me with resources so I could undertake the research described above. This work has been very close to my heart ever since I started working in the field of ancient DNA and population genetics and, with the help of the Gibbs Fellowship, I was finally able to realize my research plans. The research activities I carried out in 2017-2018 have given my research program in South Asia a big boost as I now have long-term collaborations in place as well as valuable data that will lend itself to publications and pilot results to support larger grant applications. In addition, I also want to thank the entire Newnham community for inviting me to be a part of the college for the year. I enjoyed interacting with other Fellows and students during lunches, dinners, and postdoc committee meetings. I really look forward to visiting and sharing my research with everyone in May. Additionally, this research would also not have been possible without the support of my Indian collaborators and the local communities who have partnered up in our scientific quest.

Some photographs taken during our sampling and outreach visits (taken with participants' permission)

Conversing with a community elder from the Nair population in Kerala about their folklore and ideas relating to origins and unique cultural traits

Collecting height and weight data; Nair population from Kerala



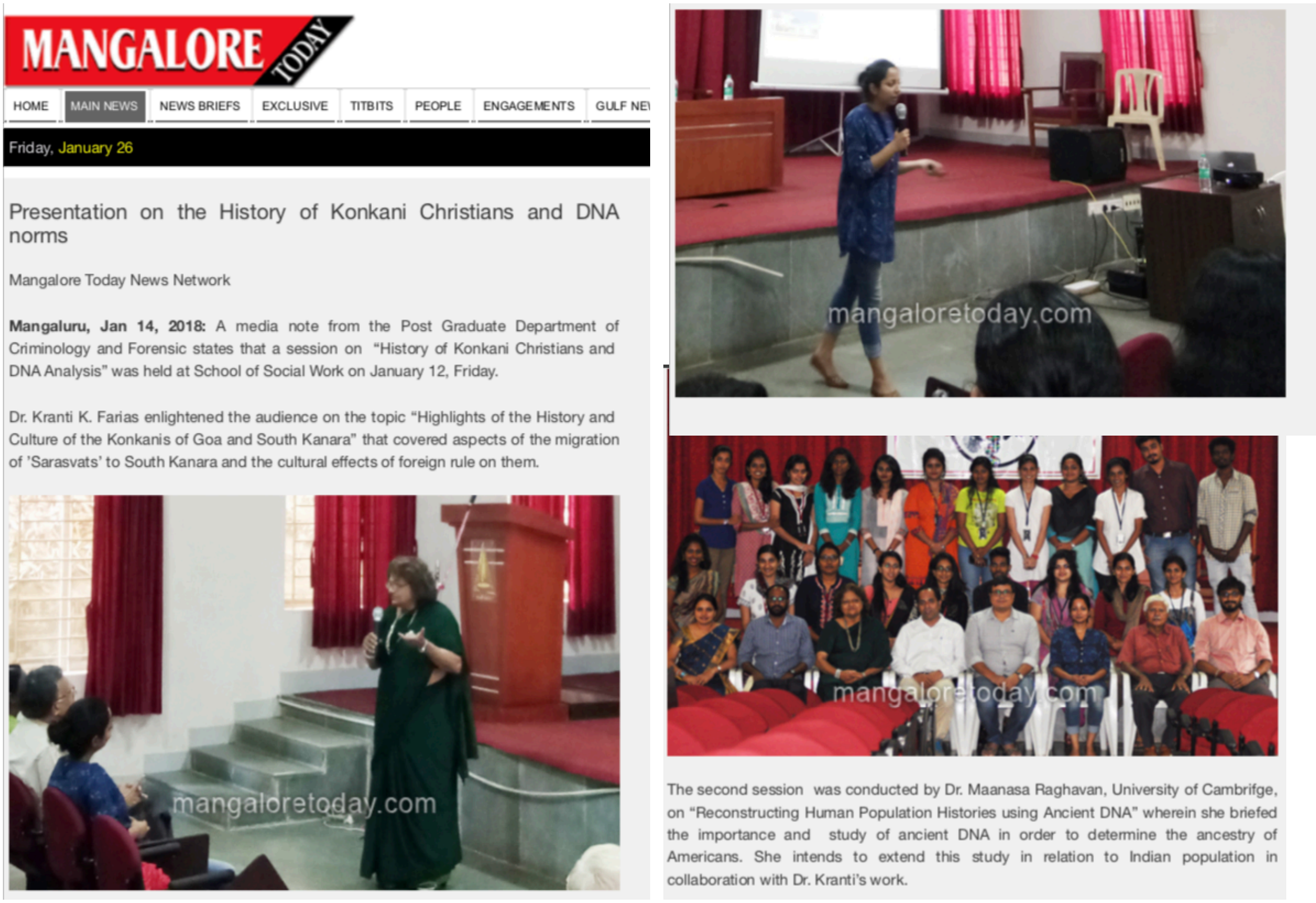
Collecting melanin index data using a spectrometer; Nair population from Kerala

Sampling team with members of the Kapla Tribe in Karnataka

Getting set up to collect blood sample from a Kapla Tribe member, Karnataka

Various stages of sampling: blood collection; database entry on laptop; phenotype data collection – Kodava population, Karnataka

Collecting melanin index data from a participant in Mangalore, Karnataka



Friday, January 26

Presentation on the History of Konkani Christians and DNA norms

Mangalore Today News Network

Mangaluru, Jan 14, 2018: A media note from the Post Graduate Department of Criminology and Forensic states that a session on "History of Konkani Christians and DNA Analysis" was held at School of Social Work on January 12, Friday.

Dr. Kranti K. Farias enlightened the audience on the topic "Highlights of the History and Culture of the Konkanis of Goa and South Kanara" that covered aspects of the migration of 'Sarasvats' to South Kanara and the cultural effects of foreign rule on them.

The second session was conducted by Dr. Maanasa Raghavan, University of Cambrifge, on "Reconstructing Human Population Histories using Ancient DNA" wherein she briefed the importance and study of ancient DNA in order to determine the ancestry of Americans. She intends to extend this study in relation to Indian population in collaboration with Dr. Kranti's work.

Excerpt from a regional newspaper, Mangalore Today, reporting on our public seminar and discussion session with the community in Mangalore, Karnataka; several individuals approached us after the session to collaborate and participate in the study