

Kayasthas of Bengal

Legends, Genealogies, and Genetics

LUCA PAGANI, SARMILA BOSE, QASIM AYUB, CHRIS TYLER-SMITH

A study of the legendary migration of five Brahmins, accompanied by five Kayasthas, from Kannauj in North India to Bengal to form an elite subgroup in the caste hierarchy of Bengal, combines genetic analysis with a reappraisal of historical and genealogical works. This combination of historical and genetic analysis creates a new research tool for assessing the evolution of social identities through migration across regions, and points to the potential for interdisciplinary research that combines the humanities and genetic science.

In the early 20th century, a debate erupted among the Bengali intelligentsia in India over the historicity of genealogical literature, which claimed that the Bengali King Adisur had invited five Brahmins from Kannauj, an ancient city in the northern Gangetic plains located in the present Indian state of Uttar Pradesh, to migrate to Bengal, in eastern India.¹ According to legend, these five Brahmins from Kannauj were accompanied by five Kayasthas, who became an “elite” subgroup described as “kulin” among the Kayasthas of Bengal.²

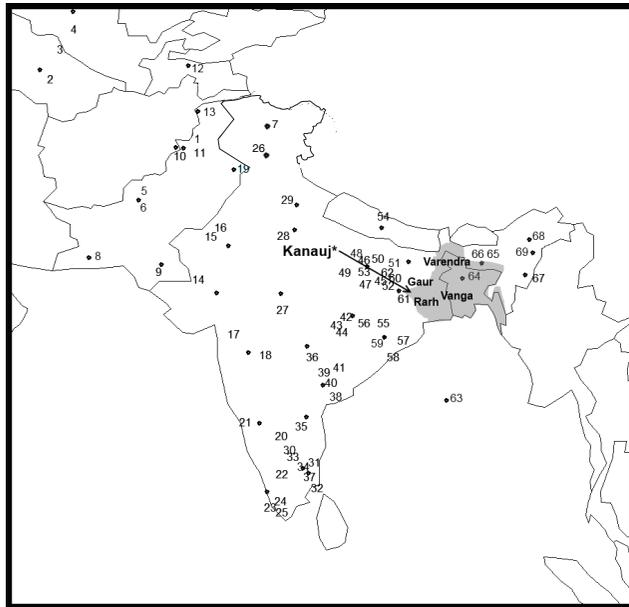
Hindu communities labelled “Kayastha” are found all over northern India, but historically, their social ranking was not uniform. At different times and in different places, those labelled Kayastha were accorded the same status as Brahmins, Kshatriyas or Sudras, and there was even a claim that they formed a fifth varna within the Hindu caste structure.³ In the popular legend, King Adisur is portrayed as the founder of “kulin-ism” in Bengal, a system of social ranking which accorded some lineages a special higher status within the Brahmin and Kayastha varna.

This study explores the legend of the migration of kulin Kayasthas from Kannauj to Bengal by combining a reappraisal of historical and genealogical works with a genetic analysis of a small group of individuals belonging to present-day Bengali kulin Kayastha families. Genetic analysis creates a new and more powerful research tool with which to assess the evolution of these social identities through migration across regions and historical events affecting relative status. The genetic trail left within present-day Bengali Kayasthas is compared to the historical and genealogical claims asserted through other means, to enable a reappraisal of the legend of Kannauj and kulin-ism in the social hierarchy of Bengal.

This study first recounts the legend of the migration of Kayasthas from Kannauj to Bengal as depicted in historical and genealogical works, including two competing theories that claim to explain the appearance of kulin Kayasthas as an “elite” subgroup in the caste hierarchy of Bengal. It then describes the data set gathered for this study, the limitations of both the genealogical and genetic data, and the methodologies applied for the genetic analysis. The original legend is reconsidered in the light of the genetic findings and their interpretations. We conclude with a discussion of the potential future paths of interdisciplinary research combining humanities and genetic science, to explore further the meanings of social constructs such as kulin-ism in Bengal, and the evolution of social identities in India.

Luca Pagani (lp.lucapagani@gmail.com) is a researcher at the Estonian Biocentre and at the Department of Biology, University of Padova, Italy. Sarmila Bose (sarmila.bose@politics.ox.ac.uk) is a senior research associate with the Centre for International Studies, University of Oxford. Qasim Ayub (qasim.ayub@monash.edu) was staff scientist at the Wellcome Trust Sanger Institute, UK, and is presently associate professor and director, Genomics Facility, at the School of Science, Monash University, Malaysia. Chris Tyler-Smith (cts@sanger.ac.uk) is head of the Human Evolution team at the Wellcome Trust Sanger Institute, UK.

Figure 1: Legendary Direction of Migration of Five Brahmins, Accompanied by Five Kayasthas, from Kannauj to Bengal at the Invitation of King Adisur



Points indicate sampling location and number assigned to each sampled population (see Supplementary Table 1 for population details). The Sivastava from Uttar Pradesh (#63) who was sampled in this study is shown as a dot in the middle of the Bay of Bengal for ease of interpretation. The territory of historical Bengal is shown in dark grey. This map is intended to provide a graphical example for the accompanying textual matter, is not to scale, and does not claim to conform to the external boundaries of India accurately. The same applies to all maps presented in this article.

Migration from Kannauj: The Origin Legend

The region referred to as Bengal approximates the present-day Indian state of West Bengal as well as present-day Bangladesh (Figure 1). This area comprised several distinct subregions in earlier centuries, delineated by the Rivers Ganga, Brahmaputra and Meghna, which flowed into the Bay of Bengal.⁴ The five kulin Kayastha lineages in Bengal, which according to legend are descended from the Kayasthas who accompanied the five Brahmins from Kannauj, are Bose, Ghosh, Mitra, Datta, and Guha in the region known as Rarh or southern Bengal (Figure 1). According to genealogical historian Nagendra Nath Basu, in the genealogies of Rarh, all Kayasthas except “Gautam Bose, Soukalin Ghosh, Viswamitra Mitra, Kashyap Guha and Bharadwaj Datta” (where the prefixes refer to *gotra* or clan), were considered “Gour-kayastha,” or local, indicating that these five lineages were the “migrants.” In northern Bengal the “non-locals” included “Vatsya Sinha” and “Moudgalya Das,” as well as “Kashyap Datta;” in eastern Bengal they included “Moudgalya Datta,” “Nag,” “Nath” and “Dam.”⁵

The original legend has a lack of clarity about the exact relationship between the five Brahmins and the Kayasthas who accompanied them. The kulin Kayastha lineages of Bengal contest any suggestion that they were the “servants” of the five Brahmins, which would suggest that they belonged to the labouring lower castes of northern India. Instead, they claim to belong to the scribal/administrator/warrior strata of the caste hierarchy of northern India. In a popular saying in Bengali, the Dattas reject the notion that they were anyone’s servant: “*Datta karo bhritya noy, songey eshechhe*” (Datta is nobody’s servant, but has come along with). Basu cites a variation:

“*Datta karo shishya noy*” (Datta is nobody’s disciple). He argues that the idea that the Kayasthas were “servants” is a misunderstanding of their becoming “disciples” of the different Brahmins at the Sena court in Bengal in the 11th–12th century.⁶

The story of King Adisur remains in the realm of legend, with some scholars questioning whether he existed at all.⁷ Mention of such a king is found in the *Ain-i-Akbari*, the 16th century work by Abul Fazl, adviser to the Mughal Emperor Akbar, among lists of Hindu dynasties and the years of their reign.⁸ Regardless of the authenticity of King Adisur, however, whether the five Kayasthas from Kannauj accompanied the Brahmins to Bengal, and whether these migrants were the founders of the kulin Kayastha lineages of Bengal, is also open to question.

Basu rejected as an unfounded myth the story of five kulin Kayasthas arriving from Kannauj with five Brahmins at the invitation of King Adisur and joining the existing Kayasthas in Bengal as an elite subgroup. In his view, all the Kayastha families had been living in Bengal long before King Adisur and therefore all were “local.” However, of 99 “surnames” in the *kulagranthas* (written genealogies), 12 were listed as *siddha* (blessed/elevated): **Bose, Ghosh, Guha**, Mitra, Datta, Nag, Nath, Das, **Dey**, Sen, Palit and Sinha (the ones examined in this study in bold). The remaining 87 were described as *mool Gour Kayastha* or original Bengal Kayastha: **Nandi, Pal, Indra**, Kar, Bhadra, Dhar, Aich, Sur, Dam, Bardhan, Shil, Chaki, Adhya, and so on. Basu argued that “Adisur” was actually King Bhaskaravarman of Kamarupa (Assam) who had conquered Rarh or southern Bengal—“Adi” meaning the “original” or first of his dynasty in Rarh, and “sur” meaning conqueror.⁹ Basu also argued that there was a second “Adisur, Adisur Jayanta, who invited five Kannauj Brahmins to come to Bengal in the 8th century.”¹⁰ In other words, there seems to be a consensus that five Brahmins from Kannauj came to Bengal at the invitation of a Hindu king, but disagreement about when this migration took place and whether the Brahmins were accompanied by five Kayasthas.

According to Basu, there was evidence of Brahmins conducting rituals and Kayasthas supervising royal matters long before King Adisur, from around the 5th century AD. Basu argued that the special status of some of the Kayastha lineages came about not because they arrived with the five Brahmins of Kannauj, but because the King of Bengal honoured these lineages at his court.¹¹ Basu attached particular significance to the year 1072, when Vijay Sena became King of Gaur (western Bengal) and Samal Varma the King of Vanga (eastern Bengal). According to him, many Brahmins and Kayasthas assembled for the coronations at the respective capitals, confusingly both called Vikrampur. Vijay Sena’s court was attended by Makaranda Ghosh, Kalidas Mitra, Purushottam Datta and Dasharath Bose (of the Mahinagar Bose clan), while Birat Guha attended the court in eastern Bengal.¹²

Regardless of the uncertainties about who King Adisur was and whether he even existed, there appears to be a consensus that kulin-ism was institutionalised in Bengal in the 12th century by King Ballala Sena, a member of the Senas (1097–1223), the last Hindu dynasty to rule Bengal. Before the Senas,

the Pala dynasty (750–1161) had been patrons of Buddhism, which moved away from social hierarchies based on birth. By contrast, the Senas, who migrated to Bengal from southern India (from the region in present-day Karnataka) in the 11th century, were strong devotees of Hinduism. Their institutionalisation of kulin-ism included compiling genealogies of kulin lineages of the Brahmin, Kayastha, and Baidya varnas which formed the main echelons of the upper castes in the hierarchy of the caste system in Bengal.¹³

The higher status of the kulin lineages within their varna was linked to the claim that they were descendants of immigrants from Kannauj.¹⁴ As an ancient Hindu city in northern India, Kannauj was at the heart of Brahmanical Hinduism. Migration between Kannauj and Bengal would be plausible despite the distance of nearly 700 miles, as there were historical connections between the two. King Dharma Pala of Bengal (775–812) had expanded his kingdom all the way to Kannauj from the original Pala base in northern Bengal, a region then known as Varendra. In the 11th century, Kannauj became a target of raids by Mahmud of Ghazni during his invasions of India. Muslim invasions from the north-west may have made the Brahmins and Hindu scribal elites of Kannauj more open to migration eastwards. By the end of the 12th century, Muhammad of Ghor had established himself in northern India and in 1204 one of his Turkish commanders, Muhammad bin Bakhtiar Khilji, led a cavalry charge into Bengal. King Lakshmana Sena, reportedly taken by surprise as he sat down for lunch, abandoned his capital and fled further east.

Maintaining Caste Status

The institutionalisation of kulin-ism was based on rules regulating marriage practices, influenced particularly by king Ballala Sena. Violation of the rules of matrimony led to loss of social status and even expulsion from the caste community. Marriage practices to enforce group endogamy became the central feature of kulin-ism, particularly due to political power in Bengal passing to Muslim rulers by the turn of the 13th century, as there was no Hindu king to grant honours to lineages any more.¹⁵ As this kind of group endogamy was practised over centuries, the kulin Kayasthas of Bengal may be expected to be a genetically close-knit group.

Marriage practices as the instrument to maintain caste status also affected the rest of the Kayastha community. Nirad C Chaudhuri wrote that there were regular discussions in his ancestral village in Mymensingh district about the “evil of *mésalliance*”: “... the blue blood of a Chaudhuri of Banagram was acknowledged as readily in Kishorganj and elsewhere as it was taken for granted at Banagram ... The five families of our common Banagram house were particularly proud of the fact that even among the Chaudhuris they were the only people who had never married, nor given in marriage, below them.”¹⁶ This kind of pressure to intermarry within specified caste strata suggests that the Kayasthas who were not kulin would also be expected to have little connection with the non-Kayastha lineages.

However, group endogamy was not watertight. The rules of matrimony themselves permitted certain types of dilution:

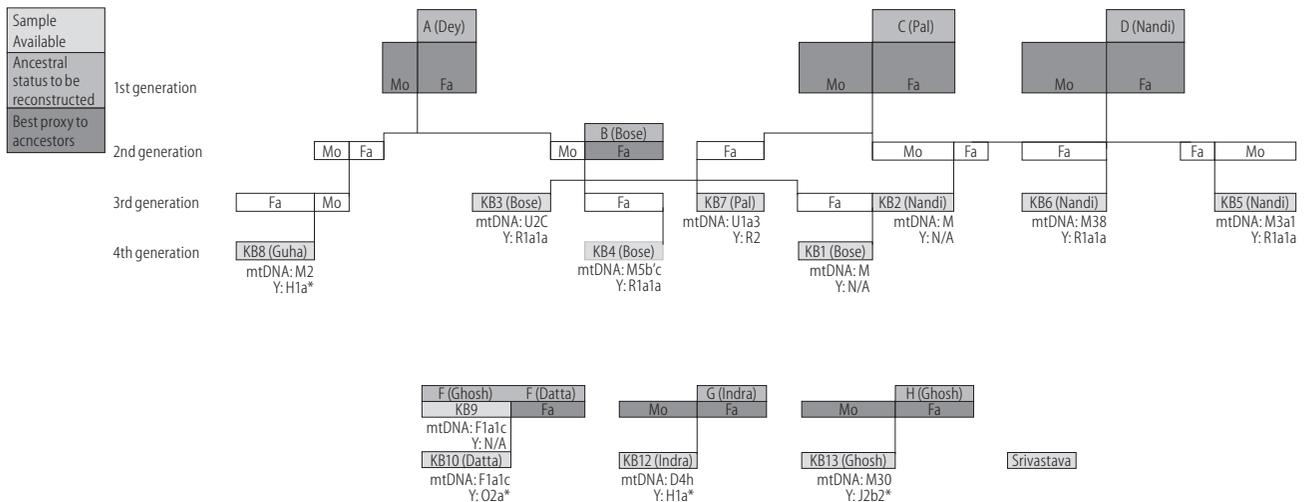
higher status Brahmin and Kayastha males were permitted to marry women from lower status families, but not vice versa.¹⁷ Moreover, it was difficult to contain liaisons within the defined caste boundaries. King Ballala Sena, who is supposed to have institutionalised kulin-ism in Bengal, was himself criticised for his relationships with women of lower castes and his alleged wish to marry a woman who was a Dom, one of the lowest castes.¹⁸

Over the centuries of Muslim rule in Bengal, many Hindus—Brahmins, Kayasthas and Baidyas—pursued successful careers, holding administrative and military positions under the Muslim sultans of Bengal.¹⁹ One of the kulin Kayastha lineages, known as the Mahinagar Bose family after the area in southern Bengal where they acquired land (*jaigirs*), held powerful positions at the court of the independent sultans of Bengal for successive generations. According to the legend, the founder of the Mahinagar Bose lineage in Bengal was Dasharath Bose, who had come to the Sena court in the 11th century.²⁰ During the sultanate period in Bengal, relations between the Mahinagar Bose family and the sultanate were so close that when the Mughal Emperor Akbar finally subjugated Bengal in the 16th century, he chose a different Kayastha family, the Pals, as his allies in the province.²¹

The most famous of the Mahinagar Boses was Gopinath Bose, known by his Islamic title Purandar Khan, who was finance minister and commander of the navy under Sultan Hussain Shah in the 15th century. He was simultaneously the leader of the Kayastha community in the region.²² Purandar Khan was influential enough to amend the rules of matrimony, allowing kulins to marry “*moulik*” (non-kulin) Kayasthas, except in the case of the eldest son.²³ The highly placed Kayasthas of Bengal became proficient in Persian (Farsi), adopted Persianised manners, and in some cases there were romantic liaisons or even marriage between upper caste Hindus and the Muslim ruling families.²⁴

In order to curb the tendency of Bengali Brahmins and Kayasthas to develop close social relations with the Muslim rulers they served, there was a periodic assembly of the caste community hosted by its leaders in which social verdicts would be delivered, based on the conduct of various families. Yet, material wealth and power influenced actual practice. Purandar Khan (Gopinath Bose) hosted such an assembly (*samikaran* or *ekjai*) in the 15th century.²⁵ In fact, Purandar Khan was so powerful that he was able to anoint two newly arrived chiefs from Rajasthan as kulin Kayasthas, establishing them in Bengal as the Datta family of Raina.²⁶

Given that King Dharma Pala of Bengal had expanded his kingdom up to Kannauj, migration between Bengal and northern India may have been a two-way traffic. Indeed, Basu speculated that the appearance of some Kayastha surnames in other parts of India may indicate travel *from* Bengal, rather than the other way around. In his hypothesis, there are connections between Bengali Kayasthas and Kayasthas in northern India, but the origins of these connections arise much earlier than the legend suggests. In a remarkably specific claim, Basu argued that the Bose lineage in Bengal (or “Vasu,”

Figure 2: Genealogical Relationships between the Anonymised Samples Gathered from the Kayasthas from Bengal

The labels below each Kayastha from Bengal (KB) sample show the mitochondrial (mtDNA) and Y-chromosomal (Y) haplogroups in males. Mo = mother; Fa = Father.

as the surname is in Sanskrit and Bengali) was the same as the Kayastha lineage Srivastava of northern India.²⁷

The caste system and kulin-ism have endured to present-day India.²⁸ Caste identity has become entrenched in Indian society despite a commitment by independent India to abolish it. It not only continues to govern social status and matrimony, but has also emerged with renewed vigour through caste-based quotas in education and jobs, and caste-based political parties, in particular across northern India. Yet it is a truism that social identities such as the caste hierarchy are a contrived human construct. A new dimension in the critical appraisal of the evolution of such social identities over the longue durée (long term), through migrations and changing fortunes of wealth and power, is now possible by combining historical enquiry with information about patterns of inheritance that might be revealed by genetic analysis.

Study Questions

This study explores the legend of the migration of the kulin Kayasthas from Kannauj to Bengal. Due to the major population rearrangements that were likely to have taken place in India in the last 1,000 years, little can be said about the timing and actual migratory path followed by these families. However, comparing the genetic ancestry in present-day descendants can help us address the following broad questions:

- (i) To what extent do the kulin Kayasthas of Bengal form a distinct genetic group, distinguishable from other Kayasthas and caste communities of present-day Bengal and India?
- (ii) Does the genetic analysis support a connection between the members of the Bose, Ghosh, Datta²⁹ and Guha families in our study (that represent four of the five lineages said to have come to southern Bengal from Kannauj) and the present-day northern Indian populations from the region around Kannauj?

Samples and Data Sets

To address these questions, we gathered genetic data from 12 adults, representing four of the five lineages said to have come to southern Bengal from Kannauj. The sampled individuals

belonged to the Mahinagar Bose lineage and their relatives by marriage, and a few other kulin and moulik Kayastha lineages of Bengal. Eleven volunteer donors of the study data were drawn from among the paternal and maternal relatives of a Bose individual and others belonging to the kulin Kayastha lineages of Bengal. An additional volunteer was a Srivastava from Uttar Pradesh (Figure 2).³⁰

All the samples had their DNA genotyped by 23andMe (www.23andme.com) for a set of ~500,000 genetic variants on the Illumina HumanOmniExpress platform. Four of the five kulin Kayastha lineages of Rarh Bengal—Bose, Ghosh, Datta and Guha—were covered by these samples, as well as the moulik Kayastha lineages Dey, Pal, Nandi, and Indra, who were believed to have resided in Bengal before the arrival of the kulins from Kannauj.³¹ The genotype data were anonymised (KB for Kayastha from Bengal followed by a code number), and merged with a set of 1,605 samples from 107 South Asian and neighbouring populations available from the literature (Behar et al 2010; Chaubey et al 2011; Metspalu et al 2011, the 1,000 Genomes Project 2015) to produce an overall data set consisting of 4,17,579 variants in 1,617 individuals.

Data Limitations

Before considering the findings of the genetic analysis in comparison to the genealogical histories and popular legends, it is important to consider the limitations of both historical and genetic data.

With regard to the genealogical histories, the indigenous production of caste histories needs to be seen in context. The interest in caste histories was strongest from around 1850 to 1930, coinciding with the rise of Indian nationalism. At the same time, the British colonial censuses were seen as potentially closing the door to upward mobility through the caste hierarchies, creating an incentive to lobby the census authorities to assign a higher status to particular groups.³² There was a “clear cultural and political purpose” behind these projects—“the preservation of Brahmanical social order.”³³ Tracing

origins back to northern India and Kannauj, the heartland of Brahminical culture, would be an important part of that project.

The interpretation of genealogical source materials also needs to weigh up the risks of bias. Basu's history of the Kayasthas is rich in genealogical data, painstakingly gathered from kulagranthas kept by families. As he laments himself, much of the family genealogies appear to be lost, making future scholars even more dependent on his works. Basu demonstrates a formidable deployment of Sanskrit texts, Bengali traditional literature, and historical works by Indian and European authors, to complement his study of kulagranthas. However, Basu's work seems hostile to Muslim rule in India and appears to assume an unquestioning allegiance to the Brahminical Hindu social order, which creates the risk that his selection and interpretation of genealogical and historical material may have been affected by these strongly held beliefs. There may be additional problems of errors arising when oral traditions were converted into written records, or differences of interpretation. There is a possibility that families may have created or amended their genealogies to suit their aspirations in the process of identity construction.

With regard to genetic analysis in discovering ancestry, there is an inevitable limitation when looking back over 1,000 years or more. Each of us has two parents, four grandparents, eight great-grandparents, and so on (2, 4, 8, 16, 32, 64, 128 ...), with the result that over about 30 generations, or around 900 years, each of us has a billion potential ancestors. In reality, as the world's current population is around 7 billion, and earlier populations were much smaller, these large numbers of potential ancestors did not exist, so the same person appeared in different places of a family tree. The number of potential ancestors of a single individual and world population size intersect at around 900 years ago—assuming an average generation time of 30 years—which means that before this time, every person was a potential ancestor of every other living person on earth. Therefore, ultimately, “the best answer to the question ‘Where did my ancestors come from?’ is ‘Everywhere.’”³⁴ This is especially true when exploring trajectories of human migration over the past 1,000 years in the Indian subcontinent, which has witnessed remarkable population movements and invasions over this period.

Human sexual reproduction dictates that when comparing individuals who have only one ancestor in common ~33 generations ago (that is, 1,000 years ago), the expected amount of shared genome inherited from this ancestor is in the order of 2^{-66} , which is close to nothing, considering that the haploid human genome comprises just over 3 billion nucleotides. However, any two present-day individuals may in practice share a considerable portion of their genome because they are closely related or because their genomes contain segments that have been around in their population of origin for a long time, and hence make them more similar to each other than to another person from a distant population.

Two particular data limitations affect our examination of the Kayasthas of Bengal, although these are likely to ease in the future with the availability of ever-increasing comparison

data. The first problem is the lack of population samples from West Bengal, in lieu of which we compare them with their geographical neighbours from eastern Bengal, present-day Bengalis in Bangladesh (BEB), who were sampled by the 1,000 Genomes Project. The BEB samples are assumed to be predominantly Muslim, reflecting the population of Bangladesh. Without information on whether the lineages sampled were originally Hindu, and if so, which caste, the BEB samples cannot be used to address questions at the family level, but do still provide the best available population from the geographical region of interest.

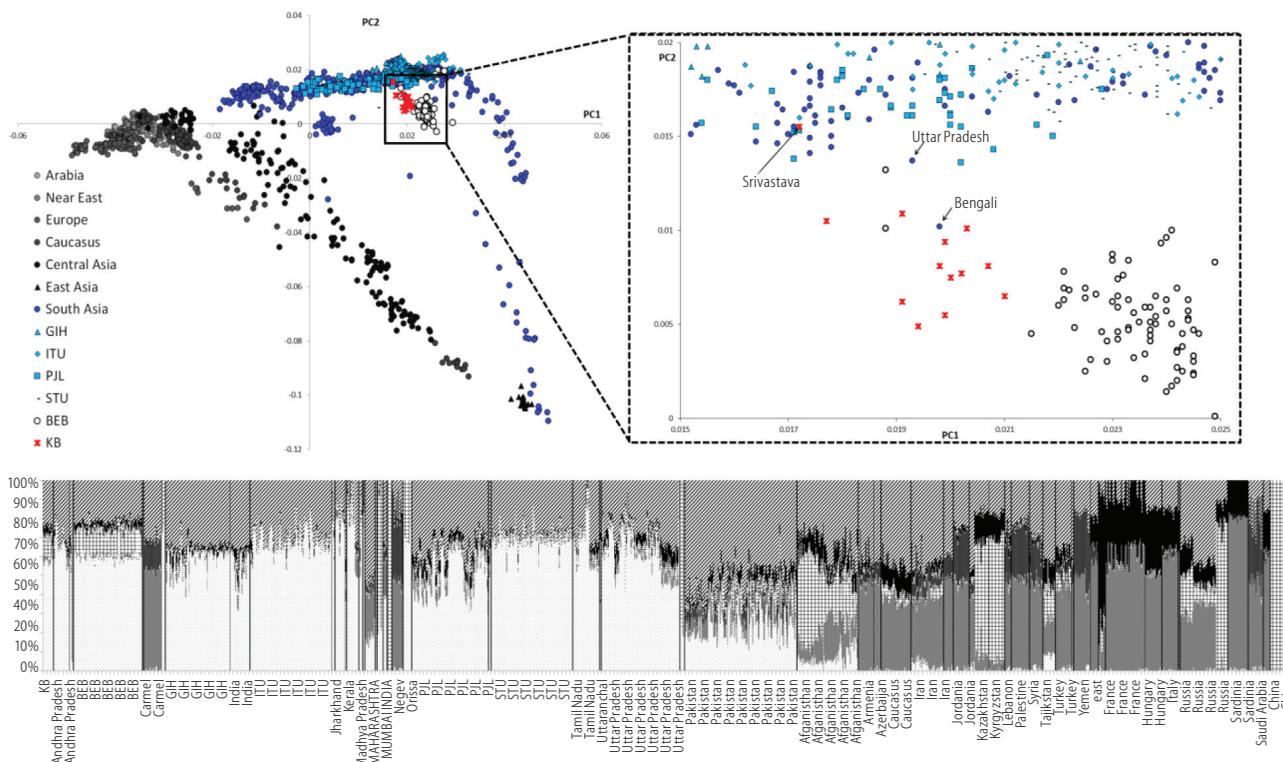
The available population comparison data also contains no samples specifically identified as “Kayastha” from Uttar Pradesh or other parts of northern India. The majority of the data comprises self-identified samples of Brahmins, a number of lower castes including “Scheduled Castes,” and many indigenous tribal communities of India. This means that the caste group of northern India to which the Bengal Kayasthas claim to be connected is largely absent. However, there are seven samples identified as “Kshatriya” from Uttar Pradesh, who may be a caste stratum similar to the Bengal Kayasthas. The other sample of an identified Kayastha lineage from Uttar Pradesh is the Srivastava sample generated and analysed in this study.

Methodology

Principal components and Admixture analyses: Principal components and Admixture are methods that we employed to infer an individual's genetic ancestry. The merged data set comprised 4,17,579 variants in 1,617 individuals (Figure 3, p 49). Admixture (Alexander et al 2009) assigns individuals to ancestral populations on the basis of variant allele frequencies at various values of κ clusters, representing the chosen number of ancestral populations. The optimum value of κ is not known in advance, and is determined from the data. In this case the best-supported model was represented by nine ancestral components ($\kappa=9$, showing the smallest cross validation error) and this model was subsequently used to infer relationships between the κ_B and other population samples. We also used the nine Admixture values to compute the Euclidean distance between each of the κ_B samples and the average values of each of the comparison populations. For each population p_i we also drew a sample s_i at random and computed the s_i - p_i Self_{*i*} distance. Each Euclidean distance between a given κ_B sample and a population p_i was then compared with the Self_{*i*} distance and, if smaller, used as the $1/\text{distance}$ radius of the circles reported in Figure 4 (p 50). Those circles are therefore to be interpreted as a measure of affinity between a given κ_B or reference sample and each of the reference populations. The bigger the circle, the higher the affinity, with simple dots representing populations for which a given sample did not yield distances smaller than the Self_{*i*} distance.

Admixture distances based on the ancestry-specific data set: To account for potential confounders arising from recent marriages between representatives of the κ_B lineages and people from outside families, we used the reconstructed genealogies

Figure 3: Principal Components (Upper Panel) and Admixture (Lower Panel, K=5) (Analyses based on 4,17,579 variants in 1,617 individuals)



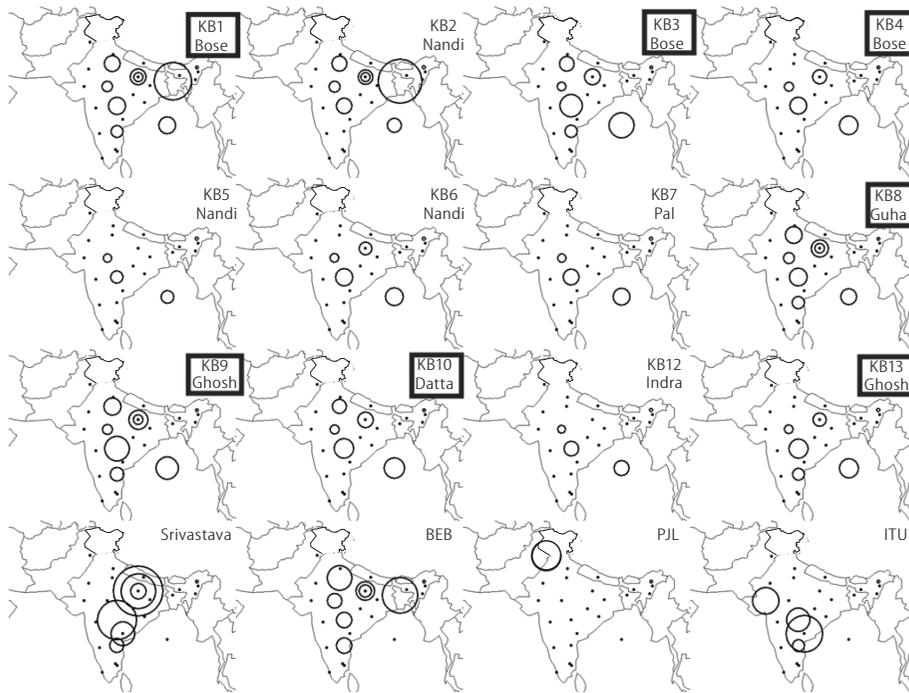
The inset shows a magnified portion of the principal components plot surrounding the Kayastha in Bengal (KB) samples (asterisks) and shows that they are close to the Bengali in Bangladesh (BEB) (white circles), their geographical neighbour in East Bengal.

reported in Figure 2 to narrow down genomic segments linked with the lineages of interest. We resolved the maternal and paternal chromosomal components by phasing our data set with SHAPEIT (Delaneau and Zagury 2012), and specifically looked for stretches of ≥ 500 variants shared between relatives. These shared genomic portions were assumed to be part of a given shared ancestry and labelled according to the scheme proposed in Figure 2. The genomic portions belonging to each ancestry were then pooled to form four “ancestry-specific”

genomes. These ancestry-specific genomes, encompassing a subset of the total variants, were then combined with the original data set, keeping only the overlapping variants. These four ancestry-specific data sets were re-run on the Admixture software and the Euclidean distances calculated as before, using $\kappa=5$ and reported in Table 1 and Figure 5. This smaller value of κ was chosen since at higher κ the ancestry-specific genomes started to accumulate their own ancestral component, becoming uninformative for the proposed analysis.

Table 1: Closest Admixture Matches of Kayastha in Bengal Ancestries with Populations Available in This Study (as shown graphically in Figure 5)

Population	Dey	Population	Bose	Population	Pal	Population	Nandi
Sakilli–Kerala	0.063	Chamar–Uttar Pradesh	0.073	Brahmins UP–Uttar Pradesh	0.084	Srivastava	0.177
Nihali–Maharashtra	0.070	Gond–Madhya Pradesh	0.079	Cochin Jews–Kerala	0.088	Nyishi–Arunachal Pradesh	0.440
Malayan–Kerala	0.094	Kol–Uttar Pradesh	0.086	PJI–Punjabi 1KG	0.095		
Paniya–Kerala	0.098	BEB–Bengali 1KG	0.098	Lambadi–Andhra Pradesh	0.099		
Gond–Madhya Pradesh	0.099	UP Low Caste–Uttar Pradesh	0.134	Brahmins KSH–Kashmir	0.103		
Pulliyar–Tamil Nadu	0.102	ITU–Telugu 1KGP	0.136	Meghawal–Rajasthan	0.108		
Mawasi–Chhattisgarh	0.119	Srivastava	0.187	Brahmins TN–Tamil Nadu	0.122		
North Kannadi–Karnataka	0.159	Nyishi–Arunachal Pradesh	0.511	Muslim–Uttar Pradesh	0.144		
Dusadh–Uttar Pradesh	0.164			GIH–Gujarati 1KGP	0.144		
STU–Tamil 1KG	0.177			TN Low Caste–Tamil Nadu	0.158		
Kol–Uttar Pradesh	0.197			Srivastava	0.164		
UP Low Caste–Uttar Pradesh	0.205			UP Low Caste–Uttar Pradesh	0.179		
ITU–Telugu 1KGP	0.208			Dusadh–Uttar Pradesh	0.195		
Chenchus–Andhra Pradesh	0.218			Chamar–Uttar Pradesh	0.241		
Cochin Jews–Kerala	0.243			Nyishi–Arunachal Pradesh	0.432		
BEB–Bengali 1KG	0.244						
Kanjars–Uttar Pradesh	0.319						
Srivastava	0.480						
Nyishi–Arunachal Pradesh	0.509						

Figure 4: Admixture Distances (K=9) between Kayastha in Bengal (KB) and Reference Samples

Each circle size is plotted as $1/\text{distance}$, where distance is the Euclidean distance computed between the admixture components of each sample and the average of a given population. Dots indicate comparisons where the sample–population distance was bigger than the distance between a population and a sample drawn at random from that population.

The circle/dot in the middle of the Bay of Bengal represents the Srivastava sample from Uttar Pradesh and was moved for ease of interpretation. Kulin surnames are boxed. South Asian populations were obtained from the 1,000 Genomes Project.

BEB = Bengali in Bangladesh; PJI = Punjabis in Lahore, Pakistan and ITU = Indian Telugu in the United Kingdom.

Comparisons with the Srivastava sample: Given the possibility of a surname shift in a branch of the KB family, leading to its formation from the Srivastava surname (or the other way around), we explicitly compared each of the available KB family samples with the Srivastava sample from Uttar Pradesh. A subset of Admixture distances ($K=9$) between KB, reference samples and a set of South Asian populations, showing only populations that are shared by both the Srivastava sample and at least one of the KB samples, is shown in Figure 6 (p 51). The geographical locations of this sample in such comparisons are shown in the middle of the Bay of Bengal.

Results and Interpretations

The studied samples—all known to be Bengal Kayasthas on their paternal side—were confirmed to be genetically closely-related and, as expected, cluster with other populations from South Asia in the principal component analysis, a widely-used method that draws population comparisons from allele frequencies. Relative to other broad groups of world populations (Figure 3) these samples cluster together and lie adjacent to BEB, their geographical neighbour in East Bengal (Figure 3–inset).

Further, their strictly paternal (Y chromosome) and maternal (mitochondrial DNA) genetic information reported in Figure 2 included, in most cases, lineages broadly diffused in the South Asian region and, hence, uninformative in tracing their specific geographical origin. An exception is provided by the Y chromosome lineage “O2a3” carried by the Kulin (Datta) KB10 sample. This lineage is characteristic of the South-East

Asian region, but has been observed at low (<5% frequency) in East Bengal and is rare (<1% frequency) in the rest of the Indian subcontinent. Such a lineage would therefore provide tentative support for a Bengal origin of this particular kulin surname. However, given the high number of generations that have occurred since the putative origin of the Datta surname, many confounding events may affect the reliability of this signal.

Overall, the Kayasthas do not show any consistent relationship to the northern Indian region of Uttar Pradesh, or with particular caste groups therein. Individuals sampled here were clearly different from each other in terms of their closeness to the population samples available for comparison, and show genetic affinity with several populations across the Indian subcontinent, including the 1,000 Genomes South Asian samples (Table 1), suggesting a history of mixed popula-

tions and a fluidity of labels of caste or community.

A word of caution is necessary about the interpretation of the matches shown in Table 1. It is important to bear in mind that the specific populations or caste groups that appear in Table 1 do not signify definitive relationships with the Kayastha samples in our study data. As explained above, the comparison data lack any samples from West Bengal or any Kayastha samples from northern India—the populations to which the Kayasthas of Bengal would be expected to be most closely related. Many other populations are also under- or un-represented. As data from the many missing populations become available, the relative position of specific groups in this figure is expected to change. What the table does show, however, is that (i) the four ancestry-specific genomes of the Bengal Kayastha lineages here are very different from each other in terms of their relationship to various available comparison population groups, and (ii) all four ancestry-specific lineages show connections to all regions of India, and not just the north Indian plains.

The diversity of the Admixture distances from different Bengal Kayastha genomic components in our study is in contrast to some other populations of South Asia. The Admixture distance analysis showed very different results for individual samples of known origin, such as Bengali in Bangladesh (BEB), Punjabis in Lahore, Pakistan (PJI) and Indian Telugu in the UK (ITU), all of which showed the best matches with other samples from their population or region (Figure 4). On this basis it is remarkable to see how the various KB samples show,

Figure 5: Admixture Distances (K=5) between Reconstructed Ancestries Focusing on Genomic Regions Shared between Individuals and a Set of South Asian Populations

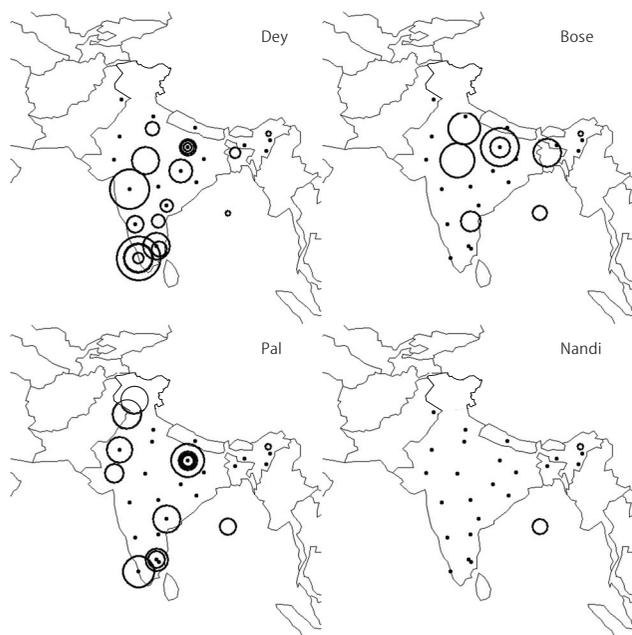
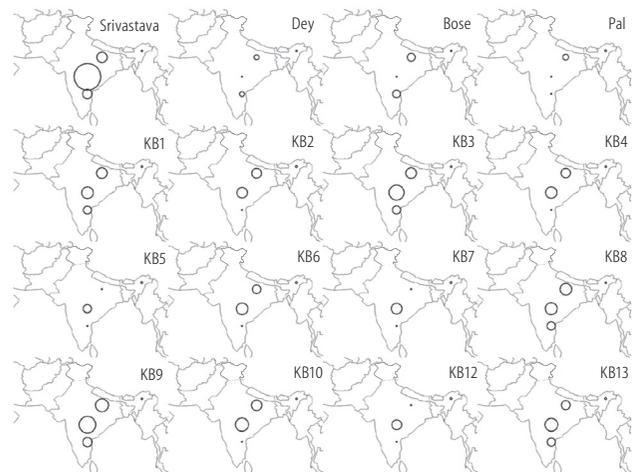


Figure 6: Subset of Admixture Distances (K=9) between Kayastha in Bengal (KB), Reference Samples and a Set of South Asian Populations



Only populations shared by the Srivastava sample and at least one of the KB samples are shown.

collectively, a variety of sharing patterns, some pointing to a marked Bengal origin (KB1 and KB2 in Figure 4), while others (KB3, KB9, KB10) show a more widespread Indian origin. Samples KB9–13 show some affinity with Srivastava and non-Bengal populations. However, we cannot determine how much of this is due to recent genetic contributions from non-kulin relatives.

In order to deconvolute the ancestral information, we focused only on genomic regions shared between individuals belonging to a given family to infer the affinity of each of the study's Kayastha samples with the available comparison populations. While Bose and Pal lineages seem to indicate a predominant origin in Uttar Pradesh and western India, the Dey

lineage points to modern south Indians as their best genetic match, while the Nandi lineage does not seem closer to any Indian population than its respective self (Table 1, Figure 5). In addition, it is worth noting how the ancestry deconvolution that was performed, for example for the Bose lineage, introduced dramatic improvements. Samples with Bose ancestry (KB1, KB3 and KB4 in Figure 4) in particular have a much less clear “kulin-like” signal than the reconstructed Bose ancestry (Figure 5). Such deconvolution was not possible for samples KB9–13 due to lack of maternal lineages.

With regard to the claim that the Srivastavas of Uttar Pradesh and the Boses of Bengal may have a common lineage, it is worth noting how the similarity between the KB data and the Srivastava sample changes depending on which study sample is considered (Figure 6). While KB3 and KB9, known to have one Bose parent, indeed show similar sharing patterns with the Srivastava sample, others (most notably KB2 and KB5 with no known Bose ancestry) do not. This could be interpreted as a distant relationship (identity) between the Srivastava and Bengal Kayastha samples, which, due to the number of generations since their separation, is currently reflected in random sharing of Srivastava lineages among the various Bengal Kayastha descendants.

Conclusions

The genetic analysis shows that the Bengal Kayastha individuals in this study are genetically closely related, regardless of whether or not they are known to be relatives in the present day. Overall, they cluster with other South Asian populations, as would be expected, particularly with their geographical neighbours (BEB) in eastern Bengal (present-day Bangladesh).

After deconvolution of their ancestral components, each Bengal Kayastha sample shows a specific signature of population sharing (Figure 5), indicating that the lineages are different from each other and highlighting their complex demographic history, suggestive of different origins and/or diverse migration paths through India to Bengal.

Do our results shed light on whether or not the legend of the migration of kulin Kayasthas from Kannauj could be true? Individuals belonging to some of the Kayastha lineages, whether termed kulin or moulik in later times, show genetic relatedness with present-day populations in Uttar Pradesh (Bose, Pal), while others show a significant genomic contribution from South India, or do not yield any informative signal on the basis of available Indian populations for comparisons (Nandi). According to Nagendra Nath Basu's definition, Pal and Nandi were reported as “local Bengal” surnames. While our genetic results seem to confirm this for Nandi (Figure 5), the variegated genetic origin of Pal points to limitations in our genetic approach, mostly due to the fluidity of labels of caste or community, and it is difficult to draw a clear distinction between kulin and moulik lineages in this regard.

Nevertheless, the high similarity to the Bose ancestry component of individuals currently living in Uttar Pradesh and North India seems to indicate a north Indian origin of this family. However, while genetic analysis might suggest a relationship, it

cannot establish directionality, nor can it say exactly when a migration may have occurred. These questions can only be answered in conjunction with historical evidence. As both the Kannauj legend and Basu's theory posited migration of the Bose lineage from north India, with Basu claiming it happened several centuries earlier, along with other Kayasthas, the genetic analysis of the Bose family provides some evidence in support of both the legend and Basu's theory in pointing to a north Indian connection, but cannot resolve their disagreement over the timing, or how the family came to be considered kulin. In addition, the conserved affinity between Bose individuals and the Srivastava sample analysed here is supportive of Basu's claim of a surname shift in the history of these two families.

With regard to potentially fruitful future research, a larger sampling of Bengal Kayasthas is unlikely to resolve the particular question of the migration of a handful of kulin Kayasthas from Kannauj at a certain time. The impact of non-kulin maternal lineages would continue to be far greater, with an increase in "noise" generated as well as any meaningful signal. More comparison samples of Kayasthas from northern India may not yield much more either, as the signals from a migration that took place 1,000 or more years ago are likely to prove too faint and dispersed to capture specific relationships.

The increasing comparison data from the Indian subcontinent as a whole and beyond, however, would continue to provide possible discovery of relationships with groups as yet unrepresented. One such avenue might be to expand the sampling to the five Brahmins whom the five Kayasthas were meant to have accompanied to Bengal according to the Kannauj legend. Such an enquiry would face the same constraints as the Kayasthas, but would generate additional data on the stories of migration to Bengal. As in this study, however, it would be important to note that despite the strong constraints introduced by the caste system, a thousand years of constant dilution of the "migrant" signature through marriage with "local" women may have irremediably erased most of the genetic signature that forms the backbone of the legend of Kannauj. Analysis of relevant ancient DNA could, in principle, overcome the complexities introduced by such dilution, but would introduce both practical challenges due to likely DNA degradation in this environment, and also potential ethical issues.

In summary, our current study has explored the application of genetic analysis to a proposed historical migration, and provided support for some aspects of it. The combination of historical and genetic research seems likely to become increasingly fruitful in the future.

NOTES

- 1 See Chatterjee (2005a). The genealogical literature comprised family genealogies referred to as *kulagranthas*, *kulajis*, and *kulapanjis*. While such genealogies had existed for centuries, Chatterjee describes the rediscovery of such material in the late 19th and early 20th century as part of a new interest in articulating a national history from the indigenous perspective in reaction to British colonial presentation of Indian history. The debate was between "rational positivists" who sought "verifiable fact" and doubted the authenticity of material like *kulagranthas*, and those who took a broader view of history and included legends, ballads, traditions, and genealogies as valuable records of social history. The latter rejected Persian and Western history as incapable of understanding Bengali society. However, Chatterjee points out that both sides agreed that it was essential to rediscover India's past through Indian agency, and while the "rational positivists" viewed oral tradition (*janasruti*) and genealogies as suspect, they did not dismiss them altogether.
- 2 The legend of the migration of five Brahmins and five Kayasthas from Kannauj to Bengal is widely recounted in popular tradition in Bengal. See Basu (1933; vol 6); Chatterjee (2005b: 173–213), and Chatterjee (2010).
- 3 Chatterjee (2010: 449). The Hindu caste structure is broadly described as comprising of four groups or varnas—Brahmin, Kshatriya, Vaishya and Sudra—but there are many sub-categories with regional variations.
- 4 Eaton (1996: 3–5).
- 5 Basu (1933; Vol 6, pp 17–29). Basu wrote several volumes of caste histories of Bengali Brahmins, Kayasthas, and Baidyas. He collected hundreds of family genealogies, and his works contain lengthy quotations from Sanskrit and Bengali texts, and elaborate family trees. While valuable for its rich sources, Basu's interpretations need to be tempered by the context of the project of indigenous rediscovery of the past, which in his case included an unquestioning acceptance of the Hindu caste hierarchy as a form of social organisation, and a marked antipathy towards Muslim rule in India.
- 6 Basu (1933; Vol 6, ch 2).
- 7 Chatterjee (2005a: 1459).
- 8 The *Ain-i-Akbari* lists several dynasties of Hindu kings by name and the number of years of their reign, starting with "Bhagrat" (presumably the Sanskrit/Bengali name "Bhagirath"), whose dynasty is described as being of the Khatri caste. Bhagrat is described as having come to Delhi on account of his friendship with "Raja Jarjodhan" and being killed in the battles described in the epic *Mahabharata*. Jarjodhan is taken to mean Duryodhan, the king leading the Kaurava side in the battle between the Kauravas and the Pandavas in the *Mahabharata*. Subsequent Hindu dynasties are described as "Kayeth" (Kayastha) and Adisur's reign is listed as 75 years. Some of the other periods of rule by individual kings exceed a century, suggesting that they may refer to multiple kings or different units of time (*Ain-i-Akbari* 1891). The *Riaz-u-s-Salatin*, a history of Bengal in Persian, written by Ghulam Husain Salim in the late 18th century, cites *Ain-i-Akbari* in its discussion of Adisur. Adisur is described as a Kayastha king whose dynasty ruled Bengal for 714 years, followed by the Pala and Sena dynasties, before the Muslim conquest by Muhammad Bakhtiar Khilji in 1198 (Maulavi Abdus Salam 1902: 50–52). Richard Eaton gives the date of Bakhtiar Khilji's conquest of Bengal as 1204 (Eaton 1996: 21).
- 9 Basu (1933; Vol 6, ch 2).
- 10 Basu (1933; Vol 6, ch 3, pp 37–38).
- 11 Basu (1933; Vol 6, ch 1). Basu argued that stone edicts and copper plates indicated that Kayasthas with names such as Mitra, Datta, Das, Nandi, Pal, Bhadra, Bose, and Ghosh were courtiers during the Gupta, Deva, and Varma dynasties in the 5th century. In one version of the story, 10 Kayastha lineages attended the court, not five.
- 12 Basu (1933; Vol 6, ch 3, p 50).
- 13 Chatterjee (2005a: 1457); Eaton (1996: 14). Lists of caste communities mentioned by the medieval Bengali poet Mukundaram described a hierarchy of four tiers, with the Brahmins, Kayasthas, and Baidyas in the first tier (Eaton 1996: 103).
- 14 Chatterjee (2005b: 176). Kannauj, also known as Kanyakubja, was an ancient city in the Gangetic plains of northern India and is now a small provincial town in Uttar Pradesh.
- 15 Eaton (1996: 103); Chatterjee (2005b: 176). Muslim rule in Bengal was heralded by a surprise cavalry attack on Nadia, the Sena capital, by Muhammad bin Bakhtiar Khilji, a Turkish commander under the Afghan invader of India, Muhammad of Ghor.
- 16 Chaudhuri (1951: 54). "Chaudhuri" is a title, usually associated with landownership. The "family name" of the Chaudhuris of Mymensingh was "Nandi," one of the Kayastha names listed by Nagendra Nath Basu.
- 17 Chatterjee (2005b: 180).
- 18 Chatterjee (2005b: 185); Basu (1933; Vol 6, ch 5, p 77).
- 19 Eaton (1996: 102).
- 20 Dasharath Bose is named by Basu as one of the Kayasthas who attended the coronation of Vijay Sena in 1072 (Basu 1933; Vol 6, ch 3, p 50). Basu lists a further eight generations of the Mahinagar Boses prior to Dasharath Bose, with the first recorded Bose in Bengal named Anantananda Bose.
- 21 Chatterjee (2010: 453–57); Basu (1933; Vol 6, ch 7, p 128).
- 22 Descendants of the Mahinagar Boses were proud of the achievements of their ancestors in the sultanates of Bengal. See, for instance, Subhas Chandra Bose's unfinished autobiography, *The Indian Pilgrim* (1980). Subhas Bose acknowledges the "well-known antiquarian and historian"

- Nagendra Nath Basu for much of the family history he recounts in this chapter, citing Basu's article on Purandar Khan in *Kayastha Patrika*, Bengali monthly, Jaistha 1335. A genealogical tree of 27 generations of the Boses, starting with Dasharath Bose, is provided in Appendix I of the book. Bose does not state where he has obtained the genealogical tree, but elaborate family trees of the Mahinagar Boses are available in Basu's *Banger Jatiya Itihas* and a whole chapter is devoted to the history of the Mahinagar Bose lineage (Basu 1933: Vol 6, ch 6, pp 99–121). "Khan" was a title bestowed by the Muslim sultans. Many other surnames in Bengal, such as Chaudhuri, Roy, Sarkar and Majumdar, are also titles, with the original family name sometimes used together with the title (for example, Datta Roy) and sometimes dropped. By the British period, the Mahinagar Boses appear no longer to use the title Khan.
- 23 Bose (1980: 8). He thought this reform had saved the Boses from "excessive inbreeding;" Basu (1933: Vol 6, ch 6, p 106).
- 24 Chatterjee (2005b: 193); (2010: 465–66). However, the idea that Hindus and Muslims could coexist in harmony was also challenged. The eminent historian R C Majumdar argued that the case of Raja Ganesh (referred to by Nagendra Nath Basu as "Ganesh Datta Khan"), the only Hindu to have taken power in Bengal from the Muslim sultans, demonstrated that Muslim conquerors did not allow Hindus to rule as king in the areas in which they had taken control (Majumdar 1968: 187–92). Basu, despite describing Raja Ganesh as someone accepted by the Muslim rulers as one of their own in court, argued that Raja Ganesh tried to re-establish Hindu independence (Basu 1929: Vol 5, pp 86–94).
- 25 Chatterjee (2005b: 200–01); Basu (1933: Vol 6, ch 7, pp 122–28).
- 26 Chatterjee (2010: 467). Basu quoted a 400-year-old ballad by Maladhar Ghatak about the Dattas of Raina and calculated the date of this event as 1487 (Basu 1933: Vol 6, ch 6, pp 101–02).
- 27 Basu (1933: Vol 6, ch 1). For the Boses of Bengal being the same as the Srivastavas of Uttar Pradesh, see p 10, citing his own *Kayasther Varna-nirnaya*, p 174, and *Itihas*, Vol 6, ch 4, p 66. Srivastava is a well-known Kayastha name in Uttar Pradesh, sometimes masked by other titles. For instance, the families of the former Indian Prime Minister, Lal Bahadur Shastri, or the Bollywood superstar Amitabh Bachchan, are said to be Srivastavas.
- 28 Chatterjee (2010: 448).
- 29 This name can be spelled Datta, Dutta or Dutt.
- 30 Informed consent was obtained from the volunteer donors to obtain ancestry data from their samples in anonymised form. All the donors are relatives or friends of Sarmila Bose. Approval was obtained from the Department of Politics, University of Oxford, for this research as academically justified for this project, and declaration was made to the head of department that there was no conflict of interest.
- 31 Note that in the kulagranthas cited by Nagendra Nath Bose, the Das and Sinha lineages are included among the "siddha" or "elevated" lineages.
- 32 Chatterjee (2005a: 1457–58).
- 33 Chatterjee (2010: 446–48).
- 34 Jobling et al (2014: 12).
- Behar, D M et al (2010): "The Genome-wide Structure of the Jewish People," *Nature*, Vol 466, pp 238–42.
- Bose, Subhas Chandra (1980): *The Indian Pilgrim*, Netaji Collected Works, Netaji Research Bureau, Volume 1, Ch 2.
- Chatterjee, Kumkum (2005a): "The King of Controversy: History and Nation-making in Late Colonial India," *American Historical Review*, Vol 110, No 5.
- (2005b): "Communities, Kings and Chronicles: the Kulagranthas of Bengal," *Studies in History*, Vol 21, No 2.
- (2010): "Scribal Elites in Sultanate and Mughal Bengal," *Indian Economic and Social History Review*, Vol 47, No 4.
- Chaubey, G et al (2011): "Population Genetic Structure in Indian Austroasiatic Speakers: The Role of Landscape Barriers and Sex-specific Admixture," *Molecular Biology and Evolution*, Vol 28, pp 1013–24.
- Chaudhuri, Nirad C (1951): *The Autobiography of an Unknown Indian*, London: Hogarth Press.
- Delaneau, O and J F Zagury (2012): "Haplotype Inference," *Methods in Molecular Biology*, Vol 888, pp 177–96.
- Eaton, Richard M (1996): *The Rise of Islam and the Bengal Frontier, 1204–1760*, Berkeley: University of California Press.
- Jobling, M et al (2014): *Human Evolutionary Genetics*, New York and London: Garland Science.
- Majumdar, R C (1968): "A Forgotten Episode in the Medieval History of Bengal," *Annals of the Bhandarkar Oriental Research Institute*, Vol 48/49, Golden Jubilee issue, pp 187–92.
- Maulavi Abdus Salam (trans) (1902): *Riazul-s-Salatin: A History of Bengal*, Ghulam Husain Salim, Asiatic Society, Calcutta, pp 50–52.
- Metspalu, M et al (2011): "Shared and Unique Components of Human Population Structure and Genome-wide Signals of Positive Selection in South Asia," *American Journal of Human Genetics*, Vol 89, pp 731–44.
- The 1,000 Genomes Project Consortium (2015): "A Global Reference for Human Genetic Variation," *Nature*, Vol 526, pp 68–74.

REFERENCES

- Ain-i-Akbari* (1891): Trans H S Jarrett, Asiatic Society of Bengal, Calcutta, pp 144–46.
- Alexander, D H, J Novembre and K Lange (2009): "Fast Model-based Estimation of Ancestry in Unrelated Individuals," *Genome Research*, Vol 19, No 9, pp 1655–64.
- Basu, Nagendra Nath (1929, 1933): *Banger Jatiya Itihas*, Vol 5: *Uttar Rarhiya Kayastha Kanda*; Vol 6: *Dakshin Rarhiya Kayastha Kanda*, Viswakosh Press, Calcutta, Bangabda, 1336, 1340.

Economic & Political WEEKLY

Review of Urban Affairs

February 18, 2017

Safe and Inclusive Cities: Pointers towards a Research Agenda

— Amita Bhide, Anant Maringanti, Darshini Mahadevia

Resistance and Its Limits: Protesting Urban Violence in Lyari, Karachi

— Nida Kirmani

Water and Conflict in Bombay Hotel, Ahmedabad:

Urban Planning, Governance and Infrastructural Violence

— Renu Desai, Shachi Sanghvi

Ecology vs Housing and the Land Rights Movement in Guwahati

— Darshini Mahadevia, Aseem Mishra, Yogi Joseph

The Violence of Worliding: Producing Space in Neo-liberal Durban, Mumbai and Rio de Janeiro

— Himanshu Burte, Lalitha Kamath

Colonising the Slum: Changing Trajectories of State–Market Violence in Mumbai

— Amita Bhide

Urban Planning and Violence: Cause or Catalyst for Change?

— Alison Brown

For copies write to:

Circulation Manager,

Economic and Political Weekly,

320–322, A to Z Industrial Estate, Ganpatrao Kadam Marg, Lower Parel, Mumbai 400 013.

email: circulation@epw.in