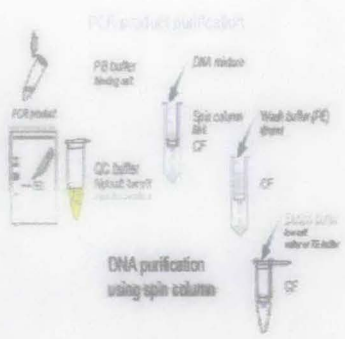
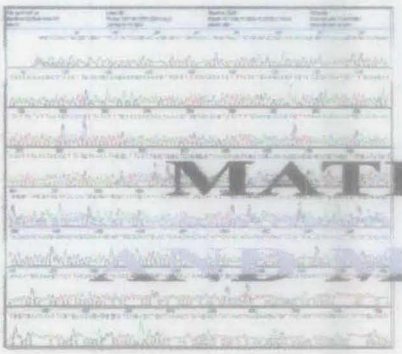
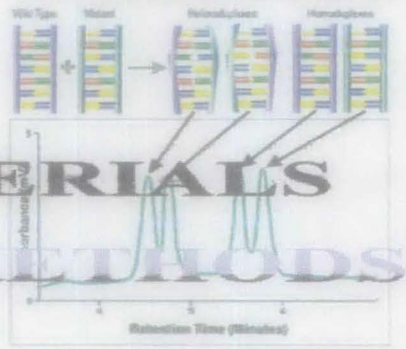
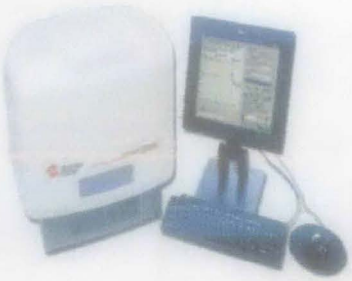
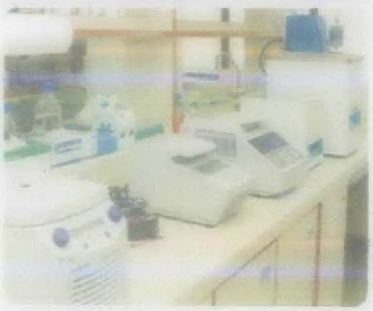
CHAPTER THREE











Materials and Methods:

Subjects/study populations:

In the present study anthropologically distinct sub-Himalayan Indian populations were chosen as they are considered as traditional populations in true sense with little or no admixture. Due to strict endogamy and socio-cultural barriers these population might have preserved their gene pool. In the present investigation some selected ethnic groups of sub-Himalayan regions of India and a few populations of Nepal and Bhutan (Mongoloid as well as Non Mongoloid populations) were considered. A total no. of 1960 samples were collected.

Subject selection:

In this present study I intensively considered samples from individuals who inhabit sub-Himalayan Indian region, Nepal and Bhutan. A short interview was conducted with every donor at the time of sample (blood) collection in this study. A set of questions was followed for interview of the donor to ensure whether their family background has link to my desired geographical study locations or not.

Brief introduction of the sampling areas:

The Himalayan Ranges

(From south to north the Himalayas are divided into different parallel ranges)

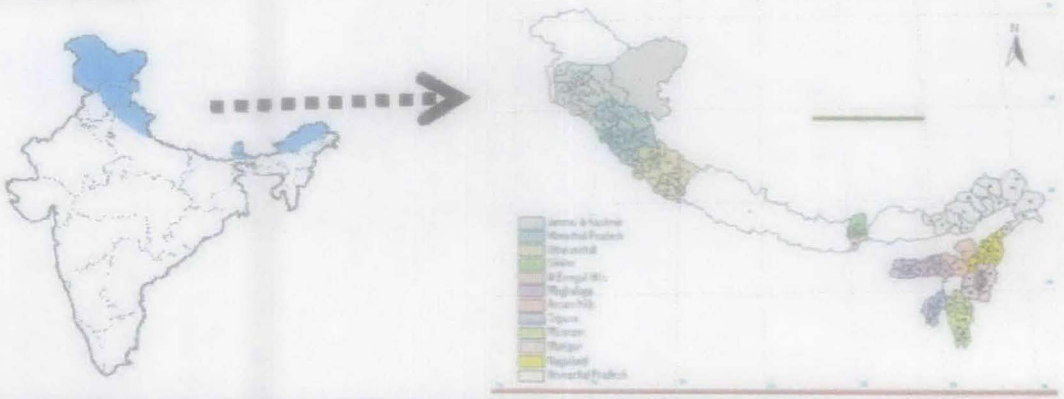
Lower Himalayas	Sub Himalayas	The Great Himalayas	Tibetan Himalayas
with an average height of 3,700 meters	with an average height of 900 to1, 200 m. It is the youngest out of the other ranges	Oldest range with a height above 6,000 m including Mount Everest, K2 and Kanchenjunga and nine of the 14 highest peaks in the world.	The Tibet Plateau is known as the Roof of the World with an average height of 4,000 to 4,5900 meters

Uniquess of the Himalayan ranges:

In terms of its diversity of peoples and languages, the Himalayan region is the most complex area of Eurasia. This remarkable landscape includes the highest land barrier on the face of the planet, and linguistic evidence shows that this has shaped and channelled population movements in the past.

History tells us that the original inhabitants of the Himalayas were the Kinnars, Kilinds, and Kiratas. Our Hindu epics and Puranas give reference of their existence in the Himalayan regions. History also mentions the names of Khasas and the Darads. But today only three different ethnic groups form the Himalayan population. They are the Negroids, Mongoloids and the Aryans. There are hundreds of different languages spoken along the length of the Himalayas. Most people speak languages belonging either to the Tibeto-Burman or Indo-European families, but there are also Austroasiatic, Dravidian, Daic and Altaic language communities settled in the mountain tracts, foothills and periphery of the Himalayas, and two language isolates, Burushaski and Kusunda.

Indian Himalaya:



Jammu and Kashmir, Himachal Pradesh, Uttaranchal, Sikkim, West Bengal hills, The North Eastern States (Seven Sister: Meghalaya,Assam, Tripura, Mizoram, Manipur, Nagaland, Arunachal Pradesh)

GENOMIC VIEW OF ETHNIC INDIA:

Indian sub-continent is a panorama of social diversities, racial differences and rich cultural heritage and highest genomic diversity after Africa.

Indian population is diverse and heterogeneous and culturally divided into tribals and non-tribals.

Regionally distinct four language families (Bhasin M.K, 2006):

Language families	Regions of India
The Austro-Asiatic Language Family (Nishada)	North & North East India
The Tibeto-Chinese Language Family (Kirata)	North East India
The Dravidian Language Family (Dravida)	South India
The Indo-European Language Family (Aryan)	Northern India

India has 22 officially recognized languages. But around 33 different languages and 2000dialects have been identified in India.

A home of 461 tribal communities they constitute 8.08% of the total population and tribal are considered to be the original inhabitants of India.

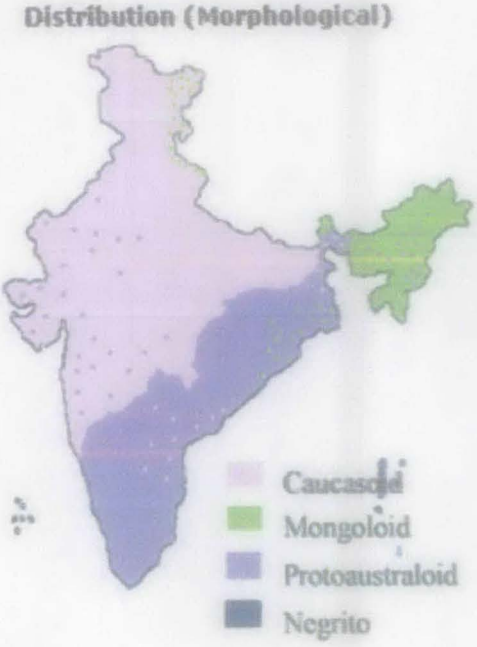
India has over 2000 castes and belongs to the Hindu religion and is hierarchically arranged in four main classes:

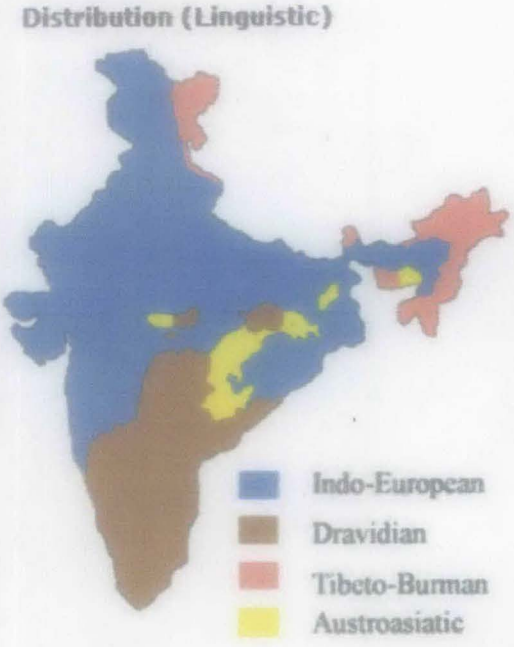
Brahmin- Priestly class> Kshatriya- Warrior class> Vysya- Business class> Sudra- Labor class

Races in India (as per classical pattern, **see Appendix** for detail classification) According to Dr. B. s. Guha. The population of India is derived from 6 main ethnic groups:

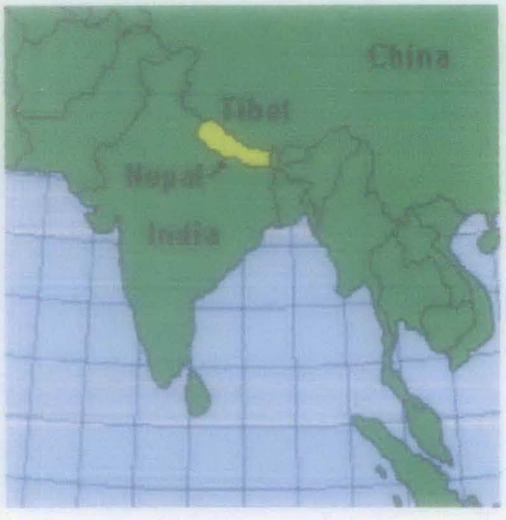
- Negrito, Proto-Australoids or Austrics, Mongololds, MediteiTanean or Dravidian, Western Brachycephals and Nordic Aryans.

Diversity of Indian Population:





GENOMIC VIEW OF NEPAL:



The people of Nepal are a complex mix of racial patterns. Nepal has a population of more than 25 million consisting of more than 70 ethnic groups having different cultures and spoken languages. The distribution of the different ethnic groups reflects the geographical diversity of the country.

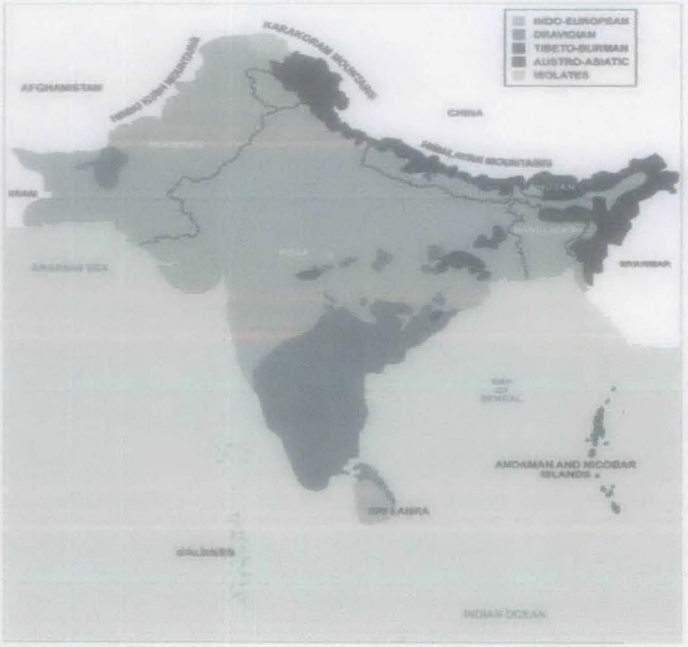
The majority of Nepal's population is of indo-Aryan origin, the remaining are of Tibetan and Bhotia,inhabitants of northern Nepal and Mongoloid inhabitants of the central belt. The major religions of the people of Nepal are Hinduism and Buddhism. Apart from these two religions the next main religion is Islam in Nepal.

GENOMIC VIEW OF BHUTAN:

There are three main ethnic, religious and linguistic groups and a dozen smaller groups in Bhutan. Bhutan is a nation of immigrants and a multi-religious, multi-cultural and multi-linguistic society. The three main ethnic groups live geographically separated with the dominant political group the Ngalong in the west, the Sarchops in the east and the southern Bhutanese of Nepali origin (also known as Lhotshampas) in the south. Bhutanese are the only large group to follow traditional Buddhi （Lamaistic Buddhist 75%）and Indian- and Nepalese-influenced Hinduism 25% of the total population(2,185,569) Dzongkha (official), Bhotes speak various Tibetan dialects, Nepalese speak various Nepalese dialects

Linguistic boundaries in South Asia:

(Ayub Q and Smith CT. ,2009. Genetic variation in South Asia: assessing the influences of geography, language and ethnicity for understanding history and disease risk, BRIEFINGS IN FUNCTIONAL GENOMICS AND PROTEOMICS. page 1 of 10 Advance Access published June 17, 2009)



Ethnographic notes on study population:

Rajbanshi:

The Rajbanshis constitute one of the largest schedule castes of West Bengal and are widespread throughout North and West Bengal and in Purnea (Bihar) and Goalpara (Assam). They are cultivating and land owning community of North Bengal, they are also known as Desi or Bahey and Rajbanshi Kshatriya. The history of the origin of the Rajbashi is a mystery. It is said that they belong to the great Bodo family that entered India in the 10^{th} Century B.C. from the east and settled on the banks of Brahmaputra and gradually spread over Assam and whole of North and West Bengal. It is generally accepted that the bulk of the populations of Eastern Bengal and Assam consisted originally of Tibeto-Burman speaking people of Indo-Chinese stock.

Rabha:

The Rabhas are one of the plains tribes of Assam. They are mostly concentrated in the districts of Kamrup, Goalpara, Dhubri, Kokrajhar, Bongagaon, Darrang and Sonitpur. The Rabha have some pocket in Mehgalaya, West Bengal, Nepal and Bangladesh.In the opinion of Grierson the term 'Rabha' is a Hindu name of the Kachari. The Rabha belong to the great Bodo group of the Mongoloid race. Major Playfair believes that after migration from the Tibetan region they settled in the Garo Hills region, and from there they spread over to different regions of Assam. Among the Rabha there are several subgroups. The Rabha are exogamous. Cross-cousin marriage is permitted but it is not considered to be preferential one. By and large the Rabha are monogamous. The primary occupation of the Rabha is agriculture. As per 1971 census, the total number of Rabhas in Assam were 1,38,630 out of which 71,497/- persons were male and 67,133 persons were females, the sex ration being 933 females per 1000 males. The Rabhas constituted 10.31% of the total tribal population of Assam, hills and plains combined.

Toto:

The Toto, a mongoloid sub-Himalayan tribe possessing a medium to high stature with stout body, is living in a partially isolated area known as Totopara, situated under the jurisdiction of Alipur Duar subdivision of Jalpaiguri district, West Bengal. The Toto have

flat noses, small eyes, broad cheeks and thick lips. Thirteen exogamous clans are evident among the Totos without any further larger division. Monogamy is the common form of marriage among the Totos but polygamy is not ruled out. They practice cross-cousin marriage. The Toto subsist on agriculture, horticulture, poultry farming, animal husbandry and pig rearing.

Mech:

Mech has maximum concentration in the sub-hilly terai regions with a population of 26959 in 1981 census. Their dialect is akin to Bodo. They are bilingual. Monogamy is the rule, plural marriage are scarce. In Mech society both sexes enjoy equal status. Among patilineal Mech household is nuclear, followed by extended type. It is said that when living beyond the pale of Hindu influence, they are omnivorous, but they did not eat flesh of elephant and cow. Though the Mech now inhabit a multi-ethnic are, thgey lead a comparatively secluded life with low social status. With the influx of the refugee, their smooth life has been disrupted. Their literacy rate is 26.97%. The younger generation has realized the importance of education and the educated youths are interested more in Jobs than in agriculture.

Lachungpa:

In Northern Sikkim the majority inhabitants, they are known as the Lachenpas or Lachungpas, meaning inhabitants of Lachen or Lachung respectively. Geographical denotations in the names of Bhutias' last names is common.The Bhutias are people of Tibetan origin, who migrated to Sikkim from various parts of India and Bhutan, some time after the 15th century. They may be migrated through the different passes ("La" in Tibetan) in the Himalayas. In Northern Sikkim for example, the Bhutias of Sikkim as a whole can be denoted as Denzongpa, or inhabitants of Denzong, the Tibetan name for Sikkim. The language spoken by the Bhutias in Sikkim is Sikkimese Bhutia language, which is 65% intelligible with either Tibetan or Dzongkha, the language of Bhutan, although in recent times Nepali is more widely spoken. Most Bhutias practice the Nyingmapa school, followed by the Kagyupa school of Tibetan Buddhism. The Bhutias are spread out over Nepal, Bhutan, and in the northern West Bengal, especially in the towns of Kalimpong and Darjeeling. Bhotey is also often used as a derogatory term, used

by people of Nepali heritage to describe people of Tibetan heritage, although most Bhutias are better off economically and educationally among the various Himalayan communities including the Nepalese.

In Sikkim, the Bhutias are mostly employed in the Government Sector, in agriculture, and increasingly in the business arena as well. In the district of Darjeeling, Bhutias are mostly employed in government offices and many are traders. Bhutias are traditionally rice eaters with animal fat fried vegetables or meat usually pork, and beef, and occasionally mutton or chicken. Other well known foods are momo- steamed meat dumplings, and the Thukpa- noodle in broth. The Losar and Losoong are two among many festivals celebrated by the Bhutia community. Almost all Bhutia festivals/holidays hold religious Buddhist significance.

Chhaang or Chyang is the favourite drink of the Bhutias, and increasingly of other communities coexisting with the Bhutias as well. It is made of fermented barley, or millets, and is served in a bamboo container called the Tongba. Tea with milk and Sugar, and butter tea is also served in religious or social occasions. Bhutias practice intermarriage within their clans and follow a very hierarchical system of bride and groom selection. Clan discrimination is widespread and marriage outside the community is looked down upon.

The ruling dynasty in Sikkim before the mid 1970s annexation by India, was a Bhutia and was from the Namgyal dynasty. (http://www.sikkimonline.info/)

Boro:

Bodos are the true ethnic Tribal Community and considered as the earliest immigrants in Assam and are said to be the most cultural community of the north-western parts of Assam. `Bodo` is derived from `Bod` that connotes Tibet, thus relating most part of the Bodos arrived from Bhutan passes. The Bodo tribal community has a belonging to a large group of ethnicity, popularly named as the Bodo-Kachari. The 6th Schedule of the Constitution of India has conferred the Bodo tribal community the status and prestige of a plain tribe.

Apart from this the Bodo tribes are also found in several places like Assam etc. Special information is being thrown about the total population of the Bodo tribal community. These Bodo tribes are amongst the primitive settlers of the valley of Assam. Kokrajhar

town is regarded as the hub of the Bodo tribal community. Other main concentration includes the Brahmaputra valleys and also its adjacent areas. Cultural exuberance of the whole of the society of the Bodo tribal community incorporates elements like dancing, singing, language etc. Also if one studies the society of the Bodo tribal community, one can find several surnames like Bargayary, Basumatary, Bodosa, Boro, Brahma, Bwiswmuthiary, Dwimary, Goyary, Ishlary, Ishwary, Khaklary, Mushahary, Narzary, Owary, Sargwary, Sibigry and Wary. As far as the languages are concerned, these Bodo tribes prefer the beautiful Bodo language, which has been derived from famous Tibeto-Burmese language family. For writing, majority of these Bodo tribes of the primitive ages make use of the Roman script and Assamese script. As a recent phenomenon, this Bodo tribal community has taken up the Nagari script. In the early years, this Bodo tribal community has practiced of all types of cultivation and farming. Rice farming, tea plantation, pig and poultry farming, and silkworm rearing are quite significant amongst them. Weaving is also a popular occupation of the Bodo tribes. All the exquisite products that these Bodo tribes have created over the years have been the main force of enabling the Bodo tribal community to reach to its zenith. Several Bodo families rear their own silkworms, the cocoons of which are then spun into silk. Amongst the Bodo tribal females, weaving has gained fame and popularity. (http://www.indianetzone.com)

Mizo:

The people of Mizoram are known by the generic name of Mizo, which literally means people (mi) of the hills (zo). It is believed that people from far-flung regions made the hills of Mizoram made their home and from the hills, they derived their name. There are numerous different tribes under the general ethnic broad group of Mizo. The state of Mizoram in India is close to both the Indo-Bangladesh border and the Indo-Myanmar border. The closeness to the numerous international borders has made Mizoram a blend of various tribes that migrated from China, Myanmar (the erstwhile Burma), and the other parts of the Northeast. Some Mizo tribes may have formed a part of the people who lived in the Tao valley in northwest China. They slowly proceeded towards the border of Tibet and Myanmar and around 1700 AD, these tribes shifted to what is known as Mizoram today. The people of the Mizoram region have acquired a multiple-cultural from

multiple tribes, who came from different parts of the eastern countries in the past. The Chinese, Burmese and Tibetan cultures influenced the people of this region to form their own lifestyles and the own craftsmanship. Today, Mizos include following tribes - Ralte, Paite, Dulien, Poi, Sukte, Pankhup, Jahao, Fanai (Molienpui), Molbem, Taute, Lakher, Dalang, Mar, Khuangli, Falam (Tashous), Leillul and Tangur. The three main sub groups in Mizoram are Lushais, Pawis and Lakhers. You can easily find those influences in the clothes worn by the women of Mizoram, the patterns and style woven and also in the woodcarvings produced in there. The absence of outside influences helped the inhabitants of Mizoram to preserve their ancient cultures and traditions of craftsmanship. Thus we can find that, the women still carrying on old traditions of weaving and the men, of basket weaving. (http://www.indianetzone.com)

Naga:

The Nagas belong to the Indo-Mongoloid family. The fourteen major Naga tribes are the Angami, Ao, Chakhesang, Chang, Khemungan, Konyak, Lotha, Phom, Pochury, Rengma, Sangtam, Sema, Yimchunger and Zeliang. The Chakhesangs were earlier known as Eastern Angamis and are a combination of the Chakri, Khezha and Sangtam sub-tribes. Now the Chakhesang tribe is spilt further; Pochury's who were earlier a part of it now claim a distinct entity. Each tribe has their own languages and cultural features. The Naga's have different stories about their origin. The Angamis, Semas, Rengams and the Lotha's subscribe to the Kheza-Kenoma legend. It is said that the village had a large stone slab having magical properties. Paddy spread on it to be dried doubled in quantity by evening. The three sons of the couple who owned the stone used it by rotation. One day there was a quarrel between the sons as to whose turn it was. The couple, fearing bloodshed, set fire to the stone which as a result cracked. It is believed that the spirit in the stone went to heaven and the stone lost its miraculous properties. The three sons thereafter left Kheza-Kenoma, went in different directions and became the forefathers of the Angami, Sema and the Lotha tribes. According to another legend, to which the western Angamis subscribe, the first man evolved from a lake called Themiakelku zie near Khonoma. The Rengmas believe that until recently they and Lothas formed one tribe. The Aos and the Phoms trace their origin to the Lungterok (six stones) on the Chongliemdi hill. Some people believe that these Indo-Mongoloids are

'kiratas' frequently mentioned in the old Sanskrit literature of which 'Nagas' were a sub-tribe. The hill tribes in the areas now known as Nagaland had no generic term applicable to the whole race. The word 'Naga' was given to these hill tribes by the plains people. This proved to be a great unifying force to the tribes now classified as Naga. Nagas are of sub-medium height, the facial index is very low, the nasal index corresponds to a medium nose, the hair is generally straight, and the skin is brownish yellow. The eyes significantly do not show Mongolian form.

Dhimal:

The Dhimals are a non-Aryan tribe of Darjeeling district and Nepal Terai. The Dhimals living in the areas of Hinduised, non-Aryan tribe and racially they belong to Mongoloid group.This primitive community seems to have disappeared from census records, mainly after the census operation of 1931, they are not mentioned distinctively. Dhimal is a little known ethnic group found only in the districts of Darjeeling. Although it is stated that in earlier times they were distributed in Andaman and Nicobar islands but at present they are found only in North Bengal only. Prof. Chatterjee and Paul Bane-Doickot include Dhimal language as the complex pronominalired of the Tibet-Burma branch and included it in the Kirat offshoot. Chenjong (1966) agrees that the "Dhimal" is an offshoot of the 'Kirat'. It is also mentioned in the census report of 1931 that the Dhimals belonged to the Tibeto-Burmese family in case of language. In this community it is found that the clan restrictions are not strictly followed. Both clan-endogamy and exogamy are followed in their community.

Kol:

The Kols are the earliest inhabitatnts of the central Himalayan region of Garwal, Kumaon and Himachal. D.N.Majumdar accredited this race with the development of Neolithic culture in India. These people are comparatively dark complexioned. There is no unanimity of views among the scholars regarding the existence of these people in hilly regions during ancient times. According to Mian Goverdhan Singh, Kols were the descendents of the Neolithic man in Himachal Pradesh. On the fringe of the Indus valley culture, the tribes inhabiting the Himalayan foothills were the Kols (Proto Australoid). Historical evidence suggests that these people migrated to the hill regions

ne Indian plains after their defeat at the hands of the more powerful Dravidian ›ple. They had to abandon their hearths and take resort to the remote areas and thick forests of the Central Himalaya.

Kharia:

The primitive aborigines of India contain variety of tribes, which shows its richness. Kharia, being an indigenous group falls in "Austroliod" group but distinctly resembles Mongoloid. The language spoken by them is also known as 'Kharia'. The arrival of Kharias belongs to 'Kolarian tribe'. Their folk songs are fully informative about their past dynamic life. Kharias are found in Jharkhand (Ranchi, Gumla, Singhbum and Simdega district), Orissa (Birmitrapur, Sundargarh, Puri, Jhunmur, Mayurbhanj), West Bengal (Mednipur) and Assam. Rislay divides six types of Kharias but actually it is parted into three. Dudh Khairias, Delki Kharias and Pahari Kharias or Hill Kharias. The Patrician family system is found among them. The name of the family is carried according to patrilocal residence and patrilineal descent. Among Dudh Kharias nine clans can be found, namely Kerketta, Bilung, Soreng, Toppo, Baa, Kiro, Kullu, Tere and Dungdung. The story regarding the origin of Kharia tribe goes like this: God created the earth and sky. He also created a hen from which egg came into existence. The white cover of egg became first Puran, which is the part of Mayurbhanj Hindus. The yellow part became the Royal family of "Bhunj". The first person was known as "Adhi Kharia". "Adhi Singh" named the first Ring. This event took place in the district of Panch-Peer in Bihar. Now it is known as Adhipur.

Parhaiya:

Parhaiya tribe is known for the unique lifestyle and culture. The Parhaiya is one of the scheduled tribes of the state Bihar, Jharkhand and some parts of West Bengal. These Parhaiya tribes do not have separate villages to live in. Rather they share the space with all the other tribes and maintain a cordial relation with all of them. Society of these Parhaiya tribes follows the trend of the common tribal community of India. In case of marriage, Parhaiya people practice monogamy. Bigamy is also permissible. However Levirate marriage and premarital relations within the lineage group are not allowed. Marriage might also take place between boys and girls of two lineages but in general it is not permitted. Generally, they pursue village exogamy. General way of getting bride is

done in lieu of paying price to the bride and also through the approval of parents of both the bride and groom. Also instances of the marriage by exchange, elopement service and love may also be cited. Birth is regarded as very joyous occasion in the society of the Parhaiya. It makes the couple fertile and washes the strain of barrenness forever. It enhances the status of the husband and the wife as father and mother. Just like any other tribal communities of Indian Territory the main occupation of these Parahiya tribal communities is mainly cultivation, including lac farming. Quite a number of Parhaiya tribes are there who survive on the resources that these Parhaiya tribes have gathered from the dense forests. People of the Parhaiya tribes, especially the women, carry on the practice of weaving beautiful baskets and clothes. That these Parhaiya are very much religious minded and this can be revealed from the fact they have developed belief on gods and goddesses. Worship of spirits, mainly, the ancestral spirits, are quite popular amongst the pious Parhaiya tribes. For better controlling and administration, these Parhaiya tribes have adapted the typical system of Panchayat. The head of the village of these Parhaiya tribes also is in the top rank in the whole system and is better known as `Mahato`. The name of his assistant is called Khato and also all the other people of each and every people of Parhaiya family are the members. (http://www.indianetzone.com)

Santhal:

The Santhals are the third largest tribal community in India after the Gond and the Bhil. According to 1981 census, the population of Santhal is 21,54,063 in the state of Bihar. They are also distributed in Bangladesh, Bhutan, Nepal and other foreign countries. In Bihar, they are mainly concentrated in Chotanagpur plateau, Santhal Parganas, and the adjacent districts of West Bengal. They migrated to other parts of India mainly as unskilled labourers of coal mines, tea gardens, brick fields, in search of their livelihood and settled there with their families. North Bengal region of West Bengal is best known for tea production. In more than 500 tea gardens, Santhals are the most predomint labourers. However, the exact number of Santhals in this area is not known. They communicate with family members by Santhali language which belongs to Austro-Asiatic family. The Santhals are endogamous and divided into exogamous clans like Soren, Hembrom, Tudu, Hansda, Kisku, Murmun, Baske, Mandi etc. The clans are gain divided into subgroups.

Oraon:

The Central and Western parts of the old Ranchi district, now split up into Gumal and Lohardaga districts in the Chotonagpur South division of Bihar, is the habitat of the bulk of the Oraon tribe. These people are scattered in several other districts of this state, Madhya Pradesh and West Bengal. The Oraon call themselves Kurukh ('Ploughmen') which is traced to one of their mythical hero-king called Karakh associated with the district of Shahabad a former home of Oraon. Linguistically, the Oraon are Dravidian and had affinities with canarese in South India.

The Oraon have totemism and number of clans. These totems are either not killed, eaten or they are not harmed. The exogamous kin ship-group is the foundation of the social system of the Oraon. They have classificatory kinship system. The marriage is monogamous. The Oraon are mainly settled agriculturist though hunting, fishing, bird catching, cattle rearing and crafts are subsidiary.

Munda:

The Munda are dispersed over Bihar, Orrisa, West Bengal, Madhya Pradesh and Tripura. The major concentration is however, in the Ranchi district of Bihar. It has more than doubled during 70 years between 1901-1981. They are settled agriculturists and live in villages. The Munda don't live in isolation. According to Munda tradition they migrated from the southern part of North India and entered Chotanagpur through modern Rohil Khand and Oudh. Schimdt regarded it as part of the Austro-Asiatic languages which is itself a stream of Austric language family which extends from Madagascar, India, Malaysia, Indonesia, Polynesia and touched the shores of South America. The tribe is divided into a number of clans known as Kili which are exogamous. All the members of a clan are considered to be descendents of a common ancestor. The Munda generally practice clan exogamy and tribal endogamy. Monogamy is the usual practice among the Munda. Instances of polygamous marriages are also found in some villages.

Muslim population:

Islam began in Arabia in the beginning of the 7th century. The founder of this religion was Prophet Muhammad who was born in 570 A.D. in a distinguished family of Mecca.

After Prophet Muhammad's death (A.D. 632) the leadership passed to Calipus or *Khalifas* who were both religious and political heads. Arabs spread Islam from the Atlantic to Sind within eighty years of Prophet's death. Though Islam proclaimed the idea of equality, but in India it has been characterized by caste. The true Muslims are divided into four large families—Saiyad, Shaikh, Pathan and Moghul. Though they are referred as castes in India, they are neither castes nor tribes but are merely names given to groups of tribes supposed to be of similar blood (Blunt 1931: p. 189). Muslims are divided into two major religious, endogamous sect—*Shia* and *Sunni* and several minor groups like *Momins, Domon, Khoja, Bohra,Moplahic.*In addition there are interior and exterior castes among *Muslims*- (a) *Ashraf* or *Sharif* meaning noble or person of high extraction. It includes all undoubted descendants of foreigners and converts from higher Hindu castes. (b) *Ajlaf Alrop* - meaning "wretches", embraces all other Mohammedans, including the functional groups and low ranking converts. Traditional occupation of *Saids* is the priesthood, while the *Moghuls* and *Pathans* correspond to *Kshatriyas* of the Hindu regime. All converts are endogamous groups and many are split into further smaller endogamous groups "amongst the *Bhands, Gujjars* and *Rangrej,* Sunis and Shias do not intermarry...; Pathan and Shaikh. Muhammadan Rajputs preserve their Hindu rules of hypergamy". "In fact, most of these castes of Hindu converts preserve some trace of their former marriage system", (Blunt 1931: pp. 201-202). Certain Muslim groups also preserve Hindu commensal taboos and religious rites. The Muslims are a predominantly rural community with marked concentration in the Kashmir valley and adjacent Kargil Tehsil, Mewat, Rohilkhand and upper Doab, Ganga Delta, Malasar and the Lakshadweep. The States and Union Territories, where Muslims are predominant, are Lakshadweep (94.84 per cent) and Jammu and Kashmir (64.19 per cent). The States in which the proportion of Muslims exceeds the national average (11.35 per cent) are West Bengal (21.51 per cent), Kerala (21.25 per cent), Uttar Pradesh (15.93 per cent) and Bihar (14.13 per cent).

Bengali population:

Bengali is native to the region of eastern South Asia known as Bengal, which comprises present day Bangladesh, the Indian state of West Bengal, southern Assam- also known as Barak Valley, and part of Tripura. With nearly 230 million total speakers, Bengali is one

of the most spoken languages in the world. Bengali is the primary language spoken in Bangladesh and is the second most spoken language in India. Along with Assamese, it is geographically the most eastern of the Indo-Iranian languages and the most eastern of the Indo-European languages. (http://en.wikipedia.org)

Bhutanese:

The people of Bhutan mainly belong to two main ethnic groups the Dukpa who belongs to Tibetan and Monpa origin and Lhotsampa. And among others there are native tribal groups such as Lepcha, Toktop and Doya. The Bhutanese are Bhutiyas of Mongolian origin who refers to themselves as Drukpa-inhabitants of Druk Yul or the 'Land of the Thunder Dragon'. Apart from a few obscure areas of Nepal and Ladakh, and Spiti in India, the Bhutanese are the only large group to follow traditional Buddhism and, despite the building of roads and controlled introduction of tourism, have maintained many aspects of the culture. The three main ethnic groups live geographically separated with the dominant political group the Ngalong in the west, the Sarchops in the east and the southern Bhutanese of Nepali origin (also known as Lhotshampas) in the south. The Bhutia, who are also called Ngalops, are the largest ethnic group and make up as much as 60 percent of the population. They are the descendants of Tibetan immigrants who came southward to Bhutan from about the 9th century onward. The Bhutia are dominant in northern, central, and western Bhutan. They speak a variety of Tibetan dialects, and the most common of these, Dzongkha, is Bhutan's official language. The Bhutia's written language is identical with Tibetan, and they adhere to the Drukpa sect of Tibetan Buddhism.

Nepali:

The population of Nepal is more than 18 million, which constitutes people belonging to different caste and races living in different regions with diverse languages, culture, and dialects. The Gurungs and Magars mainly occupy the west part of the country while the Rais, Limbus and Sunuwars occupy slopes of the eastern hills. The Sherpas live in the Himalayan region. In the capital city of Kathmandu it is the newars who mainly occupy. Among the other ethnic groups there are Satar, Rajvanshis, Tharus, Yadavas, and Dhimals. The major religions of the people of Nepal are Hinduism and Buddhism. Apart from

these two religions the next main religion is Islam in Nepal. The people of Nepal are a complex mix of racial patterns. The dominant Hindu castes of Brahmin, Thakur and Chetri, along with several others speak Nepali. The Gurungs, Magars, Tamaings, Rais and Limbus form the Gurkha regiments of the British and Indian armies. These are part of the mongoloid, tribally organized groups of hill farmers who dominate the middle hills. The Sherpas of the Solo Khumbu region in the northeast of the country are among the many Bhutiya groups who speak dialects of Tibetan.

Methodology:

Samples Collection:

Blood samples (5-10 ml by vein puncture) were drawn from unrelated adult volunteers from different endogamous ethnic population as well as caste populations from east/northeastern parts of India, Nepal and Bhutan with prior informed consent.

DNA Isolation, quantification and storage:

DNA was extracted using phenol-chloroform extraction and was stored in 10 mM Tris–1 mM EDTA (TE) (pH 8.0), as stock solutions, at -20C freezer. DNA was quantified by using Eppendorf Bio Photometer.

I. Mitochondrial DNA genotyping:

PCR Amplification of the mtDNA Genome:

DNA amplification was performed on 10-20µg of template DNA and was carried out with the thermocycler (Gene Amp PCR system 9700, Applied Biosystem) usually in the total volume of 50µl reaction mixture of the following combination for single reaction.

Deionized water- 38µl
Buffer- 5µl
BSA- 2µl
dNTPs- 2µl
Primers- Forward-1µl
Reverse-1µl
Taq Polymerase- 0.25µl
Template DNA- 1µl

these two religions the next main religion is Islam in Nepal. The people of Nepal are a complex mix of racial patterns. The dominant Hindu castes of Brahmin, Thakur and Chetri, along with several others speak Nepali. The Gurungs, Magars, Tamaings, Rais and Limbus form the Gurkha regiments of the British and Indian armies. These are part of the mongoloid, tribally organized groups of hill farmers who dominate the middle hills. The Sherpas of the Solo Khumbu region in the northeast of the country are among the many Bhutiya groups who speak dialects of Tibetan.

Methodology:

Samples Collection:

Blood samples (5-10 ml by vein puncture) were drawn from unrelated adult volunteers from different endogamous ethnic population as well as caste populations from east/northeastern parts of India, Nepal and Bhutan with prior informed consent.

DNA Isolation, quantification and storage:

DNA was extracted using phenol-chloroform extraction and was stored in 10 mM Tris–1 mM EDTA (TE) (pH 8.0), as stock solutions, at -20C freezer. DNA was quantified by using Eppendorf Bio Photometer.

I. Mitochondrial DNA genotyping:

PCR Amplification of the mtDNA Genome:

DNA amplification was performed on 10-20µg of template DNA and was carried out with the thermocycler (Gene Amp PCR system 9700, Applied Biosystem) usually in the total volume of 50µl reaction mixture of the following combination for single reaction.

Deionized water- 38µl
Buffer- 5µl
BSA- 2µl
dNTPs- 2µl
Primers- Forward-1µl
Reverse-1µl
Taq Polymerase- 0.25µl
Template DNA- 1µl

For Long PCR the combination for the single reaction (25 ml reaction volume) is as follows:

Deionized water- 38µl
Buffer- 5µl
BSA- 2µl
dNTPs- 2µl
Primers- Forward-1µl
Reverse-1µl
Long Taq Polymerase- 0.25µl
Template DNA- 1µl

Amplification cycles for PCR:

Denaturation:	94°C- 2 min.	
Denaturation:	94°C- 2 min.	Repeat 40 cycles
Primer annealing:	52°C- 1min.	
DNA synsthesis:	72°C- 1min.	
Final synthesis:	72°C- 5min.	

Amplification Cycles for Long PCR:

Denaturation:	94°C- 1 min.	
Denaturation:	94°C- 30 Sec.	Repeat 40 cycles
Primer annealing:	65.6°C- 6min.	
DNA synsthesis:	72°C- 10min.	

Number of Cycles and annealing temperature depends on primer specificity and mtDNA quality (**See Appendix** for detail primer information). Due to heterogeneous quality of the used DNA many different primers were used both for PCR and sequencing to fill the gaps in the sequences.

Amplification check:

An aliquot of 2µl of PCR product was fractionated by gel electrophoresis in a 2% agarose gel containing 0.5Mµg/ml of ethidium bromide to assess the purity and size of the DNA fragments.

Purification of the PCR product:

All the PCR products were purified with spin columns (Western BioTechnologies).

mtDNA genotyping:

After being purified on spin columns the overlapping fragments were sequenced by means of PCR primers and different internal primers (Kong *et al.* 2003) and Big Dye Terminator chemistry (Applied Biosystem).

Genotyping by sequencing:

DNA sequencing involves determining the linear nucleotide order of a segment of DNA. There are several methods of sequencing, but most are based on the Sanger Method. This is an enzymatic method that synthesizes DNA *in vitro*. The synthesized DNA is complementary to the template DNA.By determining the nucleotide sequence of the synthesized DNA, we can deduce the sequence of the template DNA.

Reaction Components:

Template is single-stranded DNA that you want to sequence. It can be a PCR product, genomic DNA, or cloned fragment.

Primer is a short fragment of DNA that binds to one end of the template DNA. The primer provides specificity to the sequence reaction and also serves as the anchor to which nucleotides are added.

Deoxynucleotides (dNTPs) extend the primer, forming a DNA chain. All four nucleotides (A,T,G,C in deoxynucleotide form) are added to the sequencing reaction.

Dideoxynucleotides (ddNTPs) are another form of nucleotide that inhibits extension of the primer. Once a ddNTP has been incorporated into then DNA chain, no further nucleotides can be added.

DNA polymerase incorporates the nucleotides and dideoxynucleotides into the growing DNA chain.

Buffer is a solution that stabilizes the reagents and products in the sequencing reaction.

Procedure:

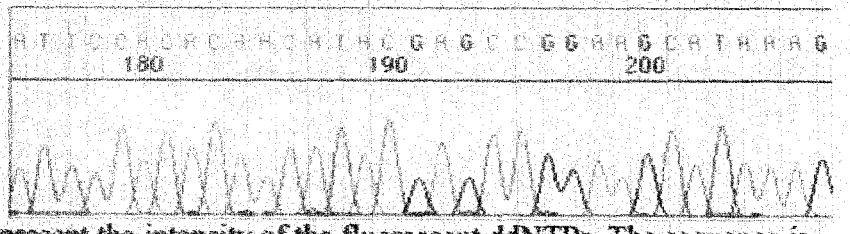
The components of the reaction are combined and allowed to incubate. Many copies of the template DNA are made by primer extension (adding nucleotides on to the primer). The copies all have the same nucleotide sequence but vary in length because the ddNTPs

incorporate randomly and stop extension. Products of the reactions are then run on a gel that separates the DNA fragments by size.

Automated sequencing:

The reaction in automated sequencing is essentially the same as in manual sequencing. There are two main differences: the labeling and reading. In automated sequencing, the products are labeled with a fluorescent dye rather then a radioactive label. There are four fluorescent dyes, each corresponding to a different ddNTP; ddATP is green, ddTTP is red, ddCTP is blue, and ddGTP is yellow. Thus each fragment has a different color at its end depending on which is the terminating nucleotide (ddNTP). This allows the products of sequencing to be run on a single lane of a gel rather then in four parallel lanes. In addition, the sequencing of nucleotides is determined by the computer, rather then being read manually by a technician. As the samples pass through the gel, a laser excites the fluorescent labels. A computer collects and analyzes this data, reading the sequence of the DNA. Thus automated sequencing is much faster and more efficient then manual sequencing. The human genome is being sequenced using automated sequencing.

This image, called an electropherogram, is the computer generated output of automated sequencing. The peaks represent the intensity of the fluorescent ddNTPs. The sequence is printed across the top of the peaks.



Assay procedure:

Purification of the PCR product:

All the PCR products were purified with spin columns (Western BioTechnologies).

Sequencing mixture:

The sequencing mixture with the total volume of 8µl consists of:

Primer- 2.5µl
DNA- 1-2µl
Big Dye- 1.9µl

The conditions for sequencing reaction:

96 °C- 10 sec.
50 °C- 5 sec. | 25 cycles
60 °C- 4min.
Hold at 4°C when complete

Purification and precipitation of sequencing PCR product:

DNA fragments attain during sequencing reaction was precipitated and purified according to the following protocol.

PCR product + 24µl Isoamyl Alcohol
↓
Shaking at 600 rpm (Left for 30 min.)
↓
Spinning at 3500 rpm for 20 min.
↓
Decant the supernatant by inverting the plates using the tissue paper
↓
Spinning at 160 rpm about 2 min.
↓
Adding 40µl Isoamyl alcohol
↓
Spinning at 3500 rpm for 7 min.
↓
Decant the supernatant by inverting the plates using tissue paper
↓
Spinning at 160rpm about 2 min.
↓
Adding of 20µl HPLC grade water
↓
Spinning at 3700 rpm for 1 min.

Analysis of the purified sequencing PCR product:

Sequencing reaction product was analyzed on a 3700 DNA Analyzer (Applied Biosystem).

Data compilation:

Sequences were handled with the DNASTAR software and aligned with the revised Cambridge Reference Sequence (cCRS [Andrews *et al.*1999]). For phylogenetic representation of data the nomenclature system of Rechards *et al.* (1998) would be

adopted. The phylogenetic trees of mtDNA would be constructed manually according to reduced median and median joining principles (Bandlet *et al.* 1999).

II. Y- Chromosome genotyping:

A total number of 78 binary markers belonging to all the major haplogroups known to be present in Indian as well as Asian populations were considered for genotyping (**See Appendix** for detail primer information). The nomenclature followed for the Y-SNP haplogroup is as recommended by the Karafet *et al,* 2008. SNP genotyping was done by three different methods:

A. Genotyping by DHPLC:

DHPLC identifies mutations and polymorphisms based on detection of heteroduplex formation between mismatched nucleotides in double-stranded PCR amplified DNA. Sequence variation creates a mixed population of heteroduplexes and homoduplexes during reannealing of wild-type and mutant DNA. When this mixed population is analysed by DHPLC under partially denaturing temperatures, the heteroduplexes elute from the column earlier than the homoduplexes because of their reduced melting temperature. Optimization of DHPLC requires the generation of a melting profile for each PCR fragment to determine its optimal temperature for analysis. Usually, the lowest temperature that shows a change in retention time of 1 minute is optimal for identification of sequence variation.

Template concentration:

In order to achieve consistent WAVE traces, it is strongly recommended that the amount of template DNA be limited to 10-100 ng. DNA extracted with some commercial systems may contain impurities that negatively influence PCR. Diluting to 10 ng may overcome these effects.

PCR amplification:

PCR amplification was performed in a 50µl reaction volume having following components: dH_2O -39.5µl, 10X PCR Buffer-5µl, dNTPs-2µl, Primers-1µleach, Taq Polymerase-0.5µl, Template DNA (10-20ng)-1µl. Following PCR conditions were applied: Initial Denaturation at 94°C for 3 mins, Denaturation at 94°C for 30sec.,

Annealing at 52-60°C for 30 sec depending on the Tm of the primers, extension at 72°C for 45 sec. This was repeated for 35 cycles and a 5 min hold at 72°C for final extension was given.

PCR buffer mix:

The PCR must be oil-free. Certain additives in polymerase buffers can have an adverse effect on the performance of the DNASep column. A list of polymerases and buffer systems recommended for the WAVE can be obtained from Transgenomics (Salisbury polymerase preferences: Optimase, AmpliTaqGold, HotStarTaq). Users are free to validate other buffer systems/ polymerases etc, but should be aware that the column warranty relating to the number of injections could be affected. Some groups reported spiking Optimase with non-proof reading Taq polymerases can overcome problems of optimization e.g. 0.05 u Platinum/1 u Optimase.

PCR product check:

Aplicons were checked by in a 2% Agarose gel before DHPLC analysis to have a quality check of the amplified product.

Post PCR processing:

This should not be necessary prior to dHPLC.

Heteroduplex formation:

I. PCR products was denatured for 5 min. at 95°C before being gradually reannealed by decreasing temperature from 95 °C to 65 °C over a period of 30 min. to enable the formation of heteroduplexes in a Gradient Thermocycler (Mastercycler Gradient, Eppendorf).

II. Each amplicon was mixed with an equal amount of known homozygous wild-type amplicon. The mixed samples were then denatured at 95 °C for 5 min. followed by a gradual decrease of temperature by 5 °C till it comes to room temperature over a period of 1 hour i.e. 5 min at 90 °C , 5 min at 85 °C , 5 min at 80 °C , 5 min at 75 °C , 5 min at 70 °C , 5 min at 65 °C , 5 min at 60 °C , 5 min at 55 °C , 5 min at 50 °C , 3 min at 45 °C , 3 min at 30 °C and 5 min at 25 °C before DHPLC analysis.

Controls:

Positive controls

If available, a mutation-positive as well as a confirmed normal DNA control should be included for each gene fragment amplified and subsequently screened on the WAVE. When a mutation control is unavailable, a confirmed normal control must be included. When amplifying large numbers of samples for many different gene fragments with a low frequency mutation pickup rate, it is sometimes acceptable to omit a normal control.

Negative controls

Negative controls containing no template DNA must be included for each amplicon (one per amplicon per plate), to check for contamination. Contamination can be excluded by agarose gel electrophoresis or WAVE analysis. It is not necessary to run negative controls on the WAVE. Asymmetric positioning of controls can help with plate orientation.

DHPLC result analysis:

It was carried out on a WAVE DNA fragment analysis system (TransgenomicTM, Crew, UK) equipped with DNASep® Column (Transgenomic TM, Crew, UK).

B. Genotyping Direct Sequencing

Following steps were followed for genotyping of Y-SNPs by genetic analyzer

I. PCR primers were selected from the published literature (Underhill *et al,* 2001; Karafet *et al,* 2008). PCR was performed in 25 µl of reaction volume having 17µl of dH_2O, 2.5µl of 10X PCR buffer, 2.0µl of 2.5 mM of dNTPs, 0.5µl of 10pm each primer, 2µl of BSA, 0.25µl of Taq polymerase. The following PCR conditions were followed: Initial denaturation at 94°C for 3 mins, denaturation at 94 ° C for 30 sec, annealing at 50-60 ° C (depending on the Tm of different primers) for 30 sec and extension at 72 ° C for 45 sec. This was repeated for 35 cycles and followed by a final extension at 72 ° C for 7 mins.

II. The amplified products were run on 2 % Agarose gel and the bands were cut from the gel.

III. DNA was purified from this gel bands by using Gel Extraction Ace Kit (Watson

Biotechnologies, Inc) following manufacturer's recommendation. The purified DNA was stored in 10 mM TE and kept in -20 C deep freezer.

IV. Cycle sequencing PCR was done in a reaction volume of 9µl having 2µl of dH_2O, 3µl of primer, 2µl of purified DNA and 2µl of sequencing kit. This was run on cycle sequencing PCR (Big Dye programme) and following PCR conditions were followed: 96° C for 10 sec, 50° C for 0.05 sec and then 60° C for 4 min and this was repeated for 25 cycles.

V. After cycle sequencing, the products were purified by employing following steps: 30µl of 75% Isoamyl alcohol was added to each well of the 96 well plate , vortexed and left for 30 mins at room temperature and then centrifuged at 3800 RPM for 30 mins. After the spin, the supernatant was discarded by centrifuging the plate upside down at 200 RPM for 1 min. 70µl of 75% Isoamyl alcohol was added again and left for 10 mins at RT. The plate was again centrifuged at 3800 RPM for 10 mins and the supernatant was discarded as mentioned above. The plate was then kept in an oven (60° C) for two hours.

VI. After drying, 20µl of dH_2O was added to the wells of the plate and sequencing was done by using 3730 ABI Sequencer.

C. Genotyping by GENOME LAB™ SNPstream Genotyping System:

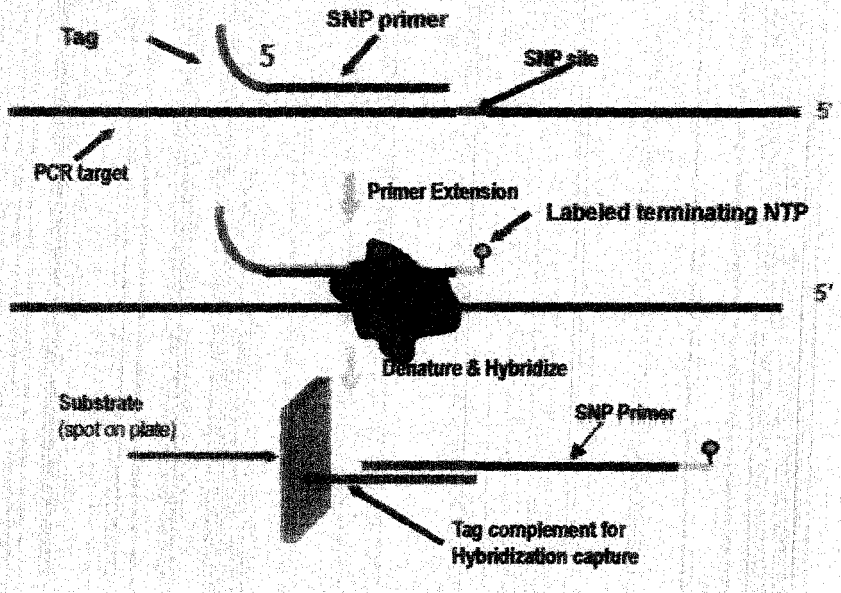
Single nucleotide polymorphism (SNP) genotyping is playing an increasing role in genome mapping, pharmacogenetic studies, and drug discovery. To date, genome-wide scans and studies involving thousands of SNPs and samples have been hampered by the lack of a system that can perform genotyping with cost-effective throughput, accuracy, and reliability.

The GenomeLab SNPstream Genotyping System provides an automated scalable system capable of performing from 4,600 to over 3,000,000 SNP genotypes per day. Combined with multiplexing capabilities, experimental layouts and operational flexibility, this enables you to process a wide range of studies.

The GenomeLab SNPstream is an automated multiplexed system that can process either 12 or 48 SNPs in each well of an arrayed 384-well plate. This system operates at a

consistent cost per sample, regardless of run throughput, allowing users to cost-effectively zero in on research targets with low throughput studies as well as conduct high-throughput analyses.

The method is based on a single-base primer extension technology in a tagged fluorescent assay. First a 12-plex or 48-plex PCR amplification is accomplished. The amplicons contain the SNPs of interest. After this multiplex amplification, the PCR products are cleaned enzimatically to remove unincorporated nucleotides and primers. In the next step two-colour fluorescent labeled ddNTPs and extension primers are added to the treated PCRs. The SNPware extension primers hybridize to the specific amplicons and the extension primers are extended by the fluorescent ddNTP, which is complement of the polymorphic site. The transfer of the labeled extension primers to the SNPware plate utilizes a tag-probe hybridization step. The extension primers are single strand DNA containing template specific sequences appended to 5' non-template specific sequences. During the hybridization step the extension primers are specifically hybridized by 5'end tag sequences to unique probes arrayed in each well of SNPware plate (**Figure 5.**).



Stringent washes remove free dye-terminators and DNA not hybridized to specific probes. The assay results are detected by direct two-color fluorescence on the SNPstream Imager. The SNPstream Imager was developed and optimized for moderate to medium high-throughput array image analysis. The Imager is based on a two-laser and two-colour approach. Each sample well is first illuminated with a 488-nm laser (blue) and with a 532 nm (green) laser beam.

Figure 5. The steps of primer extension and hybridization.

Overview of SNP Identification Technology:

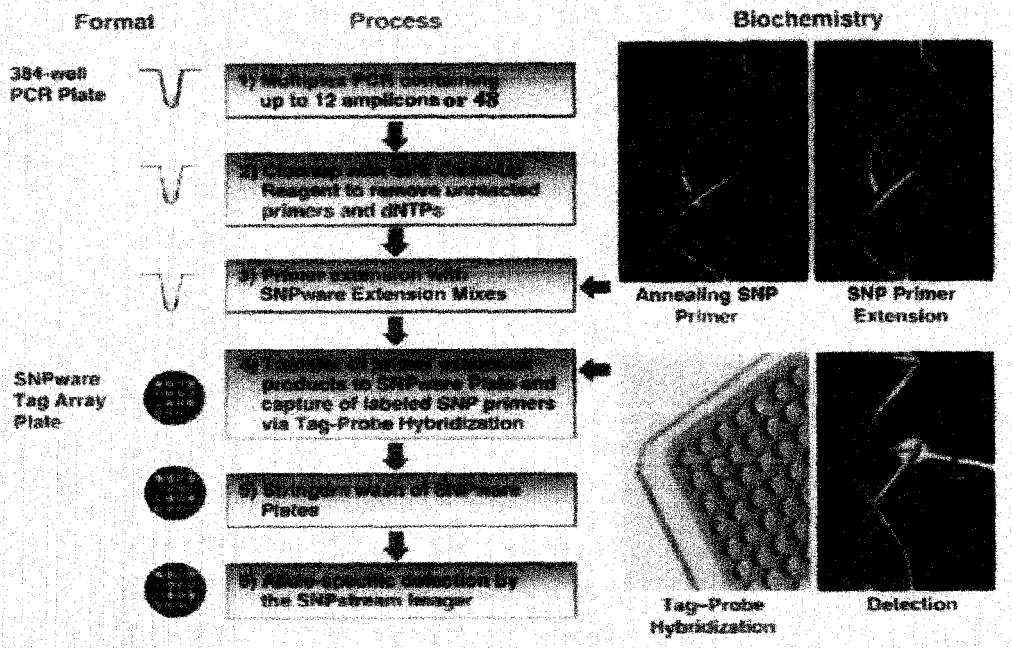


Figure 6. Process of Beckman Coulter SNPStream Genotyping Technology.

Designing of the 12plex and 48plex PCR primers

The SNP Core Facility Laboratory offers its service to design the PCR primers. Before the final SNP set is selected and primers are designed, several quality control processes have to be tested.

The quality control steps are the followings for the SNPs:

- one type SNPs (e.g.: A/G and its complement T/C) because of two-colour fluorescent detection system
- MAF > 0.05
- GC content of the sequences around SNPs between 40-65%
- repeat-free sequences around SNPs
- CNV-free sequences around SNPs
- consideration of the conserved region, haplotype blocks and the function of SNPs

After selection of SNPs we design the multiplex PCR and extension primers. The 12plex and 48plex PCR amplifications are accomplished in the same reaction condition.

Generally, the PCR-primers are ca. 20bp long and the extension primer is ca. 40bp long. Autoprimer software is used for designing of primers (www.autoprimer.com). This software was developed by Beckman Coulter Inc.

Assay Procedure:

A total number of 24 binary markers were typed by using SNPstream method. At the beginning few basic steps which are prerequisites for carrying out genotyping by this method were performed first.

These are as follows:

I. DNA was diluted to 2-10 ng/ µl.

II. PCR as well as extension primers were designed by using Autoprimer software by following the instructions of the manufacturer of the software(. Beckman Coulter, USA). A 12-PLEX System was used for genotyping. Therefore, 12 set of PCR primers were designed for multiplex PCR reaction and 12 extension primers were designed for post PCR extension of the product. While designing primers utmost care was taken for repeat masking as well as for checking GC content of the sequences. Both the quality checks were performed by using online software recommended by Beckman Coulter. Following steps were performed with the below mentioned amount of components for typing 96 samples at a time:

A. Multiplex PCR Reaction

Primer Pool (10 um each)	-2.875 µl
dNTPs (10um each)	-4.3125 µl
10X PCR Buffer II	-57.5 µl
$MgCl_2$ (25mMstock)	-115 µl
AmpliTaq Gold (5u/ul)	-11.5 µl
ddH_2O	-153.81µl
Total	-345 µl

2 µl of the DNA samples was added to each well of 96 Well Plate, followed by the addition of 3 µl of Multiplex PCR Master Mix. Plate was centrifuged for 1 min at 1000 RPM

PCR Conditions:

Initial Denaturation	94° C for 1 min	
Denaturation	94° C for 30 sec	) 40 cycles
Annealing	55° C for 30 sec	
Extension	72° C for 1 min	

Store at 4° C forever

B. Post PCR Purification

Exo I (10U/ul)	-11.5 µl
SAP (1u/ul)	-114.5 µl
10X SAP Buffer	-34.5 µl
ddH_2O	-184.5 µl
Total	-345 µl

3 µl of the purification mix was added to each well of the PCR Plate and centrifuged it briefly.

Purification Conditions:

37° C for 30 min
96° C for 10 min
4° C for forever (optional)

C. Preparation of Extension Mix

Extension Dilution Buffer	-413.7 µl
Extension Primer Mix	-3.3 µl
20X Extension Mix	-22 µl
DNA Polymerase	-2.3 µl
ddH_2O	-328.7 µl
Total	-770µl

D. Extension of the Purified Product:

PCR Conditions:

96° C for 3 min
94° C for 20 sec
40° C for 11 sec) 46 cycles
72° C for 1 min

Store at 4° C forever (optional)

E. Pre Hybridization washing of the Hybridization plate

Dilute 20X SNPware Wash Buffer I	
20X SNPware Wash Buffer I	-230.4 µl
ddH_2O	-4377.6 µl
Total	-4608 µl

14 µl of 1X SNPware Wash Buffer I was added to each well of the Hybridization Plate. The plate was reversed on a clean paper, flung on the sink and centrifuged at 150 RCF for 1 min . This step was repeated twice.

Preparation of Hybridization Solution

Hybridization Solution	-831.6 µl
Hybridization Additive	-48.4 µl
Total	-880 µl

F. Hybridization:

8 µl of the Hybridization Solution was added to each well of the PCR plate and mixed it. This was followed by the addition of 15 µl of the mixed solution from PCR plate to the Hybridization plate and left the plate in a moist chamber at 42°C for two hours.

G. Post Hybridization Washing of the Hybridization Plate

Dilute 64X SNPware Wash Buffer II	
20X Wash Buffer II	-72 µl
ddH_2O	-4536 µl
Total	- 4608 µl

14 µl 1X SNPware Wash Buffer II was added to each well of the Hybridization Plate. The plate was reversed on a clean paper, flung on the sink and centrifuged at 150RCF for 1 min. This step was repeated for two times.

The glass of the SNPware plate was cleaned by using methanol.

H. Detection

The SNPstream Imager was developed and optimized for moderate to medium high throughput array image analysis. The Imager is based on a two-laser and two-colour

approach. Each sample well is illuminated with a 488-nm laser beam and 532-nm laser beam.

I. Data analysis

The images are automatically processed into genotype calls. The genotype calls are determined by the Automated Quality Control Settings. The settings are based on the relative fluorescent intensities of two dyes and the rate of intensities. (**Figure** 7.)

Use	Description	Value
☐	Fail if Hardy Weinberg value is >	10
☐	No decision if Hardy Weinberg value is >	5
☑	Fail if Call Rate % <	50
☑	No decision if Call Rate % <	80
☑	Fail if > nn% of spots too low energy	25
☑	No decision if > nn% of spots too low energy	10
☑	Fail if YY lower relative energy >	0.14
☑	No decision if YY lower relative energy >	0.07
☑	Fail if XX upper relative energy <	0.86
☑	No decision if XX upper relative energy <	0.93
☑	Fail if XX or YY width >	0.35
☑	No decision if XX or YY width >	0.25
☑	Fail if XY width >	0.45
☑	No decision if XY width >	0.35
☑	Fail if Org Fail % >	15
☑	No decision if Org Fail % >	10
☑	Fail if lowest cluster energy < blank sample energy *	2
☑	No decision if lowest cluster energy < blank sample energy *	4
☐	Fail if XX or YY controls failed	
☐	Fail if XY control failed	
☐	Fail if XX and YY only	
☐	No decision if XX and YY only	
☐	Fail if XY only	
☐	No decision if XY only	
☐	Fail if #XY > (#XX + #YY)	
☐	No decision if #XY > (#XX + #YY)	

Figure 7. Automated Quality Control Settings

The genotype clusters can be visualized in a user-friendly scatterplot by GetGenos QCReview platform. Here we can modify the genotype clusters by manually avoiding the misclassification of samples (**Figure 8.**).

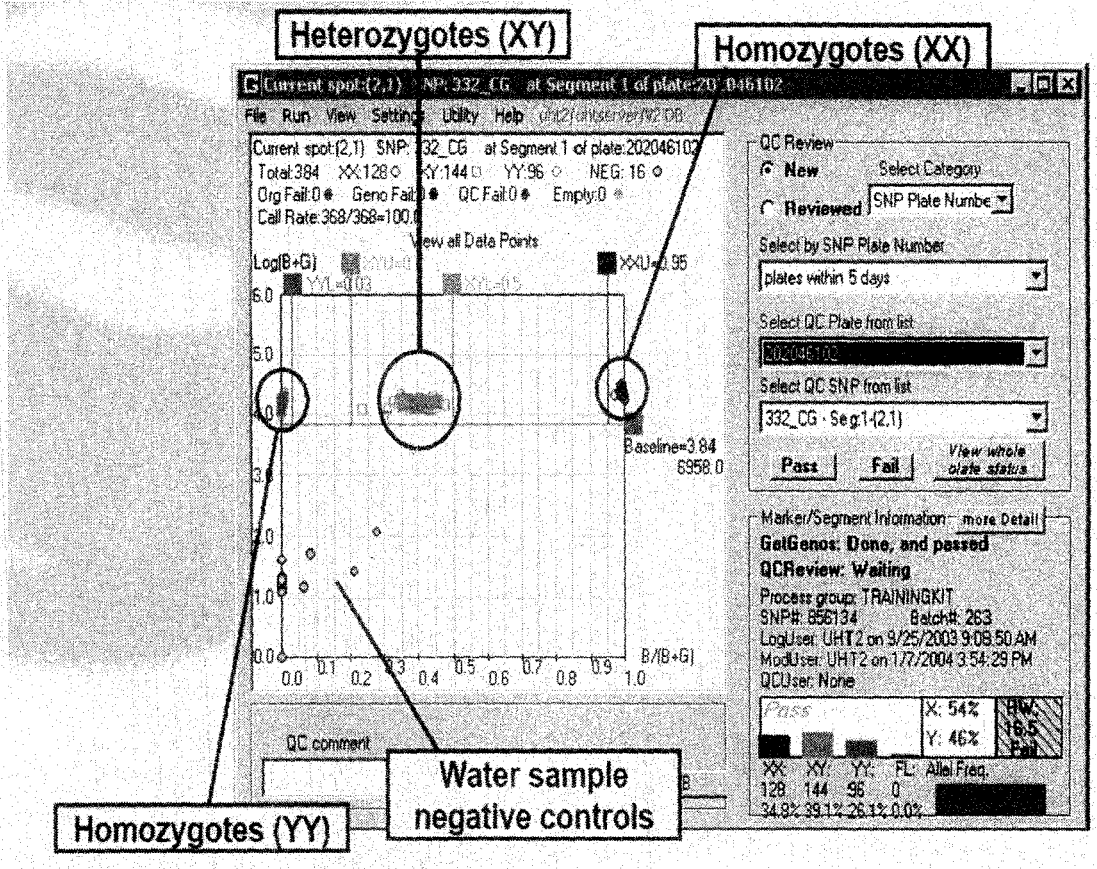


Figure 8. GetGenos QCReview .

The SNPware plate was inserted into the Genome Lab genotyping system and stepwise instructions of the software was followed for gaining the signal as well as for obtaining the raw data. At the end of analysis the raw genotype data can be exported in xls file format