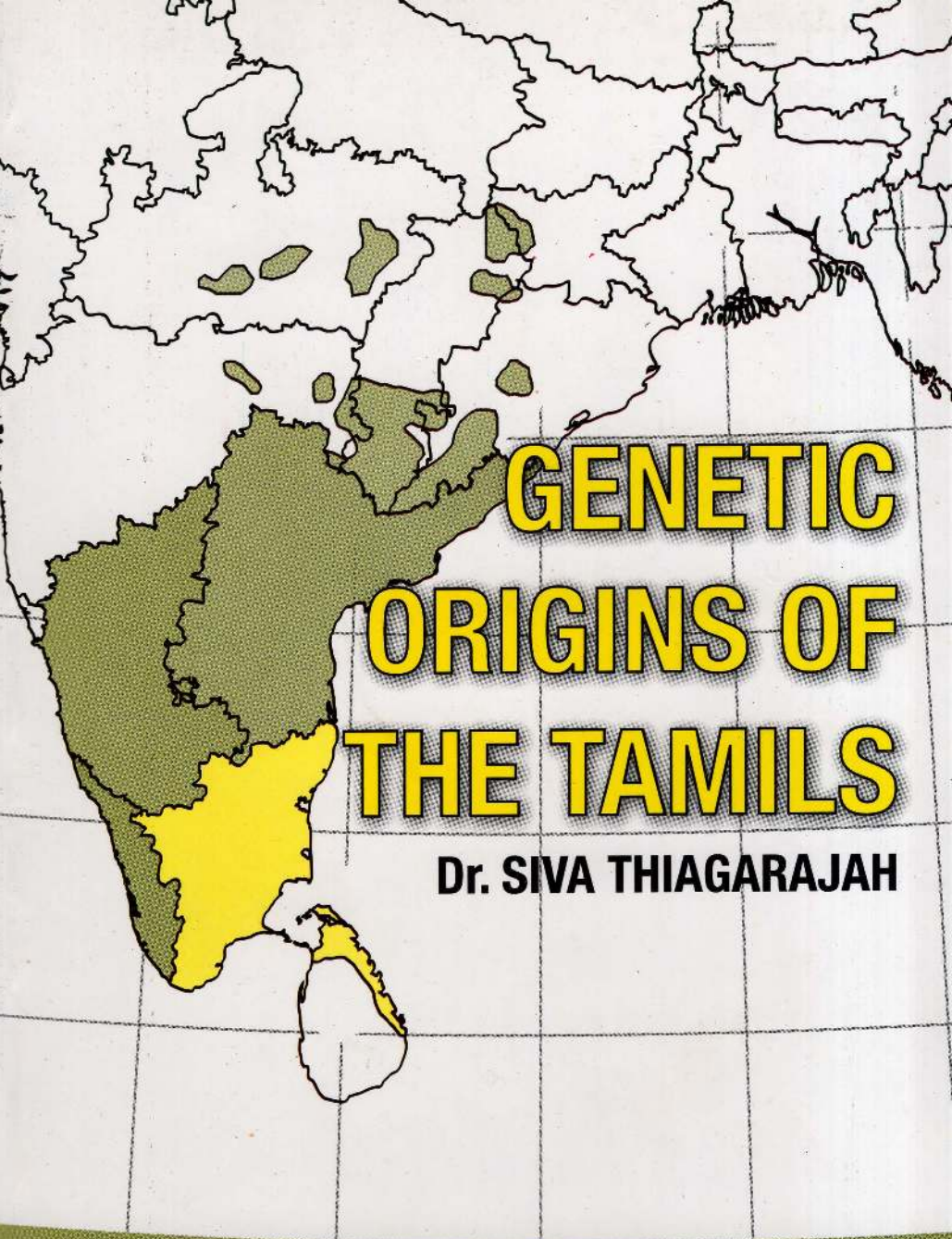


GENETIC ORIGINS OF THE TAMILS

Dr. SIVA THIAGARAJAH

Kumaran Book House

Genetic Origins of the Tamils

Genetic Origins of the Tamils

Dr. Siva Thiagarajah
B.Sc., M.B.B.S., Ph.D.

Kumaran Book House
Colombo – Chennai
2015

Genetic Origins of the Tamils
by Dr Siva Thiagarajah B.Sc., M.B.B.S., Ph.D.

Published by: Kumaran Book House
39, 36th Lane, Colombo 06, Tel.: 112364550, 113097608, E-mail: kumbhlk@gmail.com
3, Meigai Vinayagar Street, Kumaran Colony, Vadapalani, Chennai 600 026 Tel.: 2362 2680

Printed by: Kumaran Press Private Limited
39, 36th Lane, Colombo 06, Tel.: 112364550, 113097608, E-mail: kumbhlk@gmail.com

Publication No.: 660

ISBN 978-955-659-473-7

Introduction

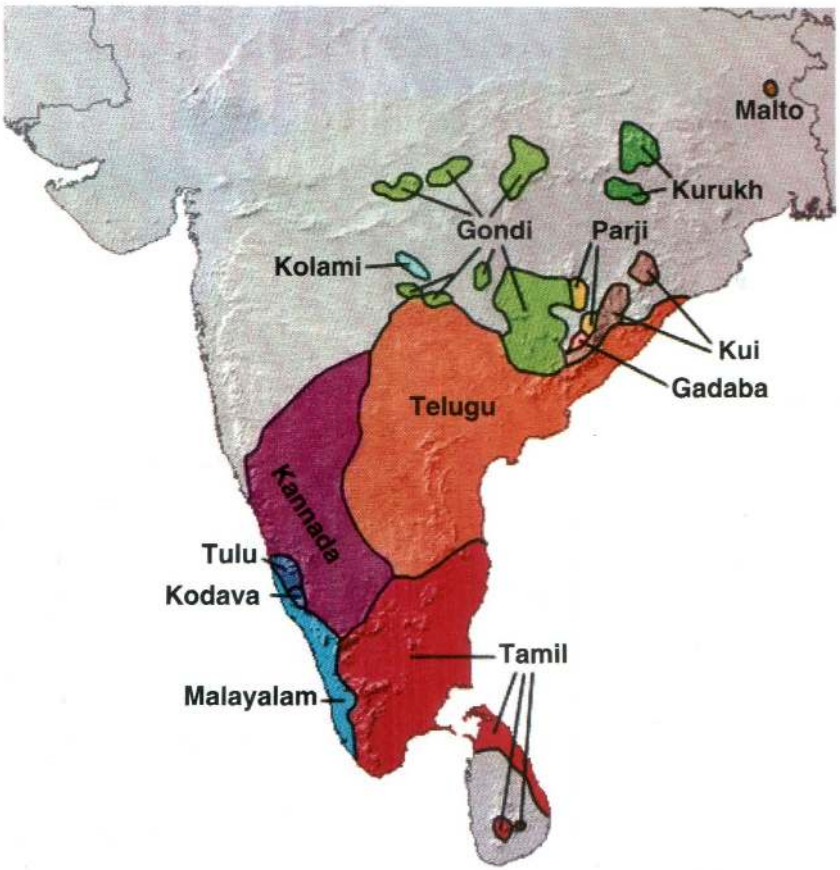
This monograph about Tamils, their genetic origins and how their language evolved from an early proto-Dravidian which have a relationship with the proto-Elamite and proto-Sumerian languages of Mesopotamia are presented here in three academic papers.

The first two papers were presented at the World Research Conference of Tamil Studies held at the School of Oriental and African Studies, London on 15 August 2013. Unfortunately the promised collection of papers presented at this conference as a book has not seen the light of the day yet. It is unlikely it will be published.

These research papers contain information about recent research findings, the views of modern day academics and the genetic findings of the past fifteen years relating to Tamils.

Dr Siva Thiagarajah

Contents



1

The Genetic Origins of Tamil Speaking People

(Paper Presented at the World Research Conference of Tamil Studies, SOAS, London 15 August 2013)

Abstract

Contrary to popular belief, there is no single Tamil race or a specific genetic or anthropological group which could be named as Tamils or even Dravidians. The Tamils are an assortment of people who arrived in South Asia over a long period of time ranging from 55,000 to 3,800 years before present. Some of them are the linear descendants of the earliest migrants from Africa to South Asia, and some of the Tamil speaking people today are of Indo-Aryan origin.

The Genetic Markers

The human body is made up of millions of living cells. Inside each cell in our body we have a nucleus containing 23 pairs of chromosomes (also called genomes). The 23 pairs of chromosomes carry in them about 30,000 genes made up of the DNA (deoxyribonucleic acid) sequences, and as they are present in pairs we have two copies of the DNA sequences. Contained within the genes are 3 billion ((3,000,000,000) building blocks known as DNA-nucleotides, made up of just four amino acids – Adenine, Guanine, Cytosine and Thymine. Among the 23 pairs of chromosomes, 22 pairs known as *autosomes* control all bodily structural and functional development while the 23rd pair, called the *sex chromosomes*, determines the formation and functioning of sexual organs. In females the sex chromosomes are denoted by XX, while in the males the sex chromosomes are denoted by XY. The Y chromosome present in one of the pair is dominant over the X.

The genes we carry in our body cells are inherited from our parents and define us as unique individuals. Genes are made up of long string like molecules – the DNA, which carries a string of coded instructions to build and maintain our bodies. Our DNA carries, hidden in its string of four simple molecules, a historical document stretching back to the origin of life-forms, the self replicating molecules, through our amoebic ancestors, to the present day. We are the end result of over a billion years of evolutionary tinkering and our genes carry the 'seams and spot-welds' that reveal this story (Spencer Wells: 2003). It is not the genetic code itself that provides the clue, but the differences we find when we compare the DNA from two or more individuals. The differences are the historical language of the genes, and it is from deciphering this genetic language we recreate our different stories.

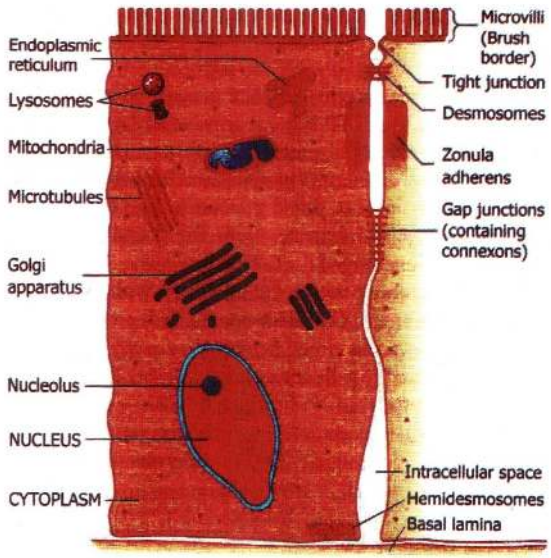
The reason we have two copies of each chromosome comes down to procreation. Only one copy of the mother's chromosome goes into each of her eggs. Similarly only one copy of the father's chromosome goes into his sperm. When the sperm fertilizes the egg, the chromosomal copy from the father and the chromosomal copy from the mother combine together, and form the new genome of the baby.

This mixing known as genetic recombination generates the diversity among the individuals. Most of the DNAs (except those in the Y-Chromosomes) get shuffled up and spliced at fertilization of the ovum at every generation. The differences in the DNA sequences occurring during splitting and recombining are called polymorphisms. Polymorphisms are found roughly once every 1,000 nucleotide of the DNA chain.

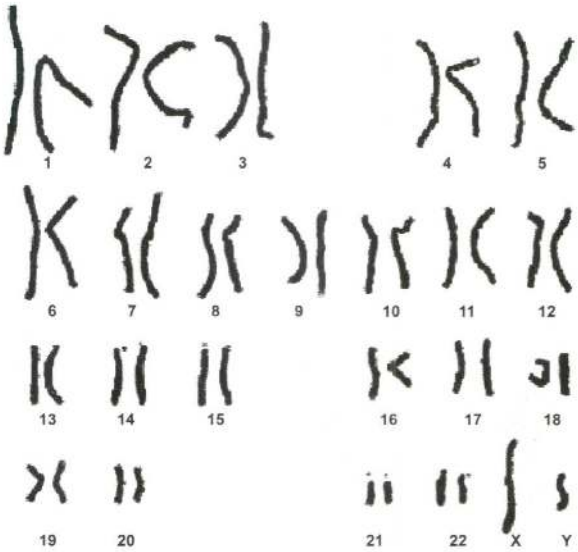
The Y-Chromosome and the DNAs contained in it is a suitable material for the purpose of population studies. This chromosome is only one of a pair, found only in the males, and passed down only from father to son. The DNAs in the Y-Chromosome do not recombine as other parts of the genome do. The Y-Chromosome is passed on as an entire chunk of DNA from father to son, basically unchanged through generations except for random changes called *mutations*.

These mutations which can happen naturally in the chromosomes are then passed on to the rest of the generations which follow them. Mutations in the Y-Chromosomes which generally occur once in 20-100 generations (roughly 500 – 2500 years) are used as the *'genetic markers.'* These mutations are harmless, but an invaluable asset to the geneticist. These markers are given a number like M8, M9, M10 etc., usually in the order in which they are discovered in a particular position in the DNA chain. These numbers are then universally recognised by the geneticists. Once a marker has been identified, geneticists can go back in time and trace it to the point at which the particular mutation – the marker – had first occurred, which would be the *most recent common ancestor.*

The point at which a common ancestor is found on the time-scale is determined by the gene we are looking at, the mutation rate, the population size and on several other factors such as whether any bottlenecks in the progression of population expansion had occurred or not.



Picture of a typical Human Cell. Mitochondria are separate organelles found outside the nucleus.



Picture showing 22 pairs of autosomes plus XY sex chromosomes present in a typical cell nucleus of a human male. Females will have XX sex chromosomes.

There is another chunk of DNA which passes through generations relatively unchanged. This is the mitochondrial DNA, which we have in our cells outside the nucleus. In the sperms of the males the head has the nucleus containing one of a pair of the nuclear chromosomes (the *haploid* chromosomes), while the mitochondria are contained in its tail. During fertilization the sperm's head enters the ovum carrying with it the spermatic chromosomes in its nucleus, while the males' mitochondria are discarded with the sperm's tail. But in the female ovum the mitochondria stay within the cell outside its nucleus, and get incorporated into the cells of the zygote (fertilized egg). Thus, the mitochondrial DNA is passed down the generations by the mothers to their children, but not by the fathers.

By studying the Y-Chromosome DNAs of the males and the mitochondrial DNAs (mt-DNA) of the females living all over the world, scientists are able to work out the genetic tree of most of the peoples inhabiting this planet Earth. Who are we? Where did our forefathers come from? When did we arrive at the land where we are living now? Which route did we take to arrive here? These and

many other questions are now being answered by the fine resolution of the genetic tree by geneticists from the United States, United Kingdom, India, China, Japan, France, Germany, Canada, Australia, New Zealand and elsewhere working on the Human Genome Project, Genographic Project and several other regional projects by individual organizations over the past two decades.

The mt-DNA study and the resolution of the genetic tree has shown that, among the multiple mt-DNAs only one was able to come out of Africa and dispersed all over the world (Cann, R.L. et al. 1987; Vigilant, L. et al. 1991; Watson, E. et al. 1997). These findings agree with the genetic trees compiled using the Y-Chromosome studies.

Haplogroups, Haplotypes and Subclades

Mutations are unique sporadic changes occurring in the nucleotide chain of a DNA referred to as Single Nucleotide Polymorphism or SNP. The Y-DNA SNP mutation has been passed from father to son over thousands of years.

Haplogroup refers to a group of people who share a series of genetic markers – which means a line of ancestors – ultimately leading to a single 'deep ancestor' in the distant past. Haplogroups are defined by a series of unique mutation events occurring as Single Nucleotide Polymorphism. These SNPs mark the branches of a haplogroup and indicate that all descendants of that haplogroup at one time shared a common ancestor. Thus, a haplogroup represents deep ancestry. It provides information about the evolutionary markers leading to the last genetic marker. e.g. Haplogroup L: *M168>M89>M9>M20*

Haplotype refers to the genetic DNA pattern you share with all males who have descended from a single forefather who had the last genetic mutation - which means your father, grandfather, great grandfather along your paternal lineage – all carry the same Y-DNA haplotype, all having the same Single Nucleotide Polymorphism (SNP) mutation. e.g. Haplotype L: *M20*

The description of a haplogroup begins with its line of descent starting from the earliest known common ancestor and moving down the line through the various branches of the family tree, listing the accumulation of genetic markers that define the haplogroup. All humans belong to a haplogroup, the ancestral clans which gave rise to their present genetic marker. So using their genetic patterns all the individuals can be sorted into particular haplogroups.

The Y-DNA SNP mutation has been passed from father to son over thousands of years. Over time, additional SNPs may occur within a haplogroup, leading to a new lineage. These new lineages are considered as Subclades of that particular haplogroup. Subclades are branches of haplogroups. Each time a new mutation occurs, there is a new branch in the haplogroup, and therefore a new subclade. e.g. Subclades of L-M20: L*, L1 (M27, M76), L2 (M317), L2*, L2a (M274), L2b (M349), L3 (M357), L3a (PK3)

In humans there are both Y-Chromosome (Y-DNA) haplogroups as well as mitochondrial (mt-DNA) haplogroups associated with the males and females respectively, which can be used to define and classify genetic populations.

The Exodus Out of Africa

During the evolutionary process of 3.8 billion years, several organisms which appeared on Earth became extinct and several others adapted themselves to the environment in which they lived and continued their existence. Those organisms which were able to adapt better, survived longer, proliferated and dominated the landscape. Among all these organisms, human beings were the most successful.

The modern man *Homo sapiens* evolved in East Africa from the African hominid (human-like primate) population between 200,000 – 180,000 years ago. He was classified as *Homo sapiens,* Latin for 'wise man' by biologist Carl Linnaeus in the 18th century. They developed a larger brain and the power of speech which the earlier hominids did not have. He was able to communicate with his fellow men using a rudimentary language, and this accomplishment made all the difference. He was able to assemble his people together,

live in groups and guard himself better against all predators and other enemies. At first he spread to all over Africa, and later left in several waves of migration to all other parts of the world, and multiplied replacing the earlier populations of *Homo erectus* in Asia and Neanderthals in Europe. Present African populations show the greatest genetic diversity, which means their ancestors are the oldest modern humans on Earth.

According to geneticist Dr. Stephen Oppenheimer, although modern human beings appeared in Africa between 200,000 and 180,000 years ago, it was only around 80,000 years ago a group of these Homo sapiens journeyed out of Africa, and their descendants spread out to the far reaches of the continents. Every human being on earth today is a descendant of this pioneer African group. This incredible journey of those pioneering people across land and mound, river and ravine, sea and ocean can now be plotted in time through a combination of archaeology, anthropology, climate study and most amazingly the study of human genes (Stringer, C. 2000; Oppenheimer, S. 2003).

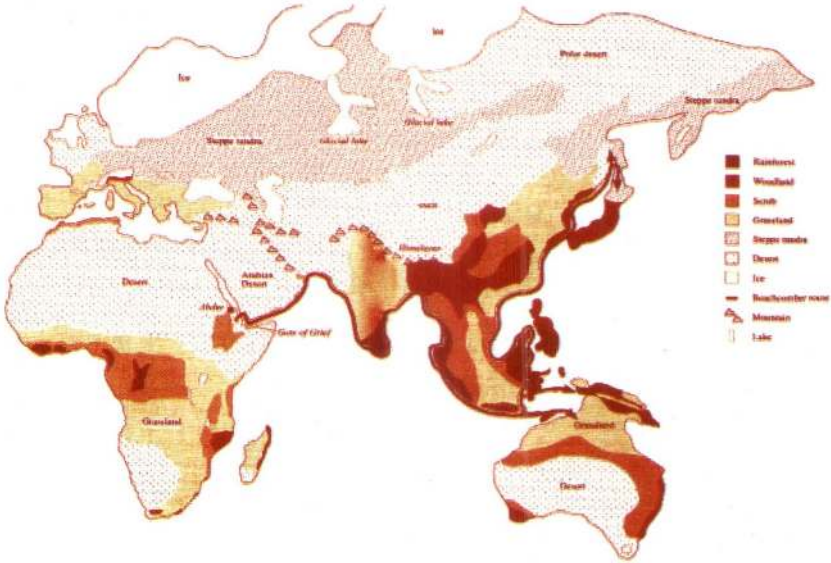
Genetic studies based on the Y-chromosomal DNA by Spencer Wells and colleagues give a more recent date of 60,000 years ago or 60,000 YBP (years before present), for the successful migration of modern man out of Africa. Spencer Wells, currently chief of the International Genographic Project has summarised his findings in his book 'The Journey of Man' (Wells, Spencer: 2002).

Based on the Y-Chromosome DNA and mt-DNA findings the route taken by these pioneer Africans is reconstructed by modern geneticists. Although there is a difference between the dates provided by the Y-Chromosome and mt-DNA studies, there is no disparity in the identification of the pioneer populations and their migratory routes.

The Y-Chromosome studies conducted by Spencer Wells and colleagues have provided consistent results over the past decade. The following is a brief summary of the migratory path taken by modern humans to inhabit the Earth, reconstructed from their studies (Wells, Spencer 2002, 2006).

Africa is physically linked to Europe and Asia via the Sinai Peninsula in the north. To come out of Africa, one has to cross the Sahara and the Sinai, then move along the Nile corridor to reach the Levant and Europe.

Another route which is closed now but available during the last glacial period when the sea levels were low was to go towards the East African coast, cross the mouth of the Red Sea, enter the Arabian Peninsula and then move towards Iran and India along the coast (Quintana-Murci, L. 1999; Petraglia, M.D. 2003). By this coastal migration, beachcombing along the south coast of Asia, the forerunners would have eaten their way to South India by 58,000 YBP (years before present). South India and its extension Sri Lanka were full of woodland and rainforest, suitable for human habitation. Hence, a significant number would have stayed behind while the forerunners marched on.



Picture showing the coastal migration of the first Homo sapiens from Africa to South Asia, South East Asia and Australia between 60,000 and 50,000 years ago. During that period of the Ice Age the sea-levels were about 100 metres below its present level when Sri Lanka was a part of the Indian Subcontinent and the Sunda Archipelago (Sumatra, Borneo, Philippines, New Guinea etc.) was a single landmass connected to South Asia.

Picture Courtesy: Stephen Oppenheimer.

From Sri Lanka they continued along the Indian Ocean coastline around Burma to Western Indonesia, then a landmass attached to Asia, and to Timor. Still following the coast they moved to South China. Paul Mellars describes this as a *'coastal express route'* with modern human populations nipping across the shores of the Indian Ocean. The genetic evidence indicates that the modern humans moved quickly eastwards, reaching Andaman Islands and Malaysia by 55,000 YBP and possibly even 65,000 years ago (Mellars, P. 2006).

From Indonesia and Timor it was easy land hops to Australia, which they reached by 55,000 YBP. Groups crossed by boat from Timor to Australia and from Borneo into New Guinea. It is important to note that there are no primates in Australia – no monkeys; no apes. Even the early hominids were never able to go to Australia. The modern man was the first human being to cross the sea and set foot in Australia! (Stringer, C. 1996; Stringer & McKie 1996; 2000; Underhill et al. 2001).

The Genetic Pool of South-Indian Tamils

The Tamils unlike some other linguistic groups do not belong to a single racial or genetic group. They originated from different genetic backgrounds but came to live together and to speak a particular language among the family of languages in South Asia forming a specific linguistic group.

Y-Chromosome DNA analysis of the South Indian Tamils conducted by several investigators over the past ten years under the Genographic Project, as well as other programmes have shown that there are about 7 major and several minor genetic groups showing specific genetic markers among the present-day Tamil speakers. The most important ones are:

1. **Haplogroup C** (Australoid group – Y Chromosome DNA marker *M130*) Evolutionary Line: *M168> M130*

The earliest carriers of this genetic marker arrived in South India about 55,000 years ago. They mark the first group of *Homo sapiens* to set

foot in South Asia. Their descendants are found among the aborigines of Australia, among the people of China, among the population of South-east Asia, Andaman islanders and the Veddas of Sri Lanka. The descendants of this marker are also present among the Pacific Islanders and among the speakers of Na-Dene languages of North America (Spencer Wells 2003, 2007).

People carrying this marker are widespread throughout the Indian Subcontinent except in the North West, where it is rare. It shows peaks in Bangla Desh, Andhra Pradesh, coastal Tamil Nadu and Sri Lanka. Although the tribal people of this genetic group still speak the Munda languages, the majority of them have mixed up with, and live among the mainstream communities and speak the regional languages. This genetic marker is present among 5% of the males in South India.

In April 2008, Prof. R.M. Pitchappan and his team from Madurai Kamaraj University identified the genetic marker *M130* among eleven members of the family of Mr. Virumandi Andithevar of Jothimanickam village in Madurai, which identified them as the direct descendants of the first *Homo sapiens* to arrive in South Asia 55,000 years ago. The Indian media gave wide coverage to this finding and this was also reported in the BBC Television in the U.K. (Thiagarajah, Siva 2011: 98).

2. **Haplogroup F**: (Middle Eastern group – Y Chromosome DNA marker *M89*): Evolutionary Line: *M168> M89*

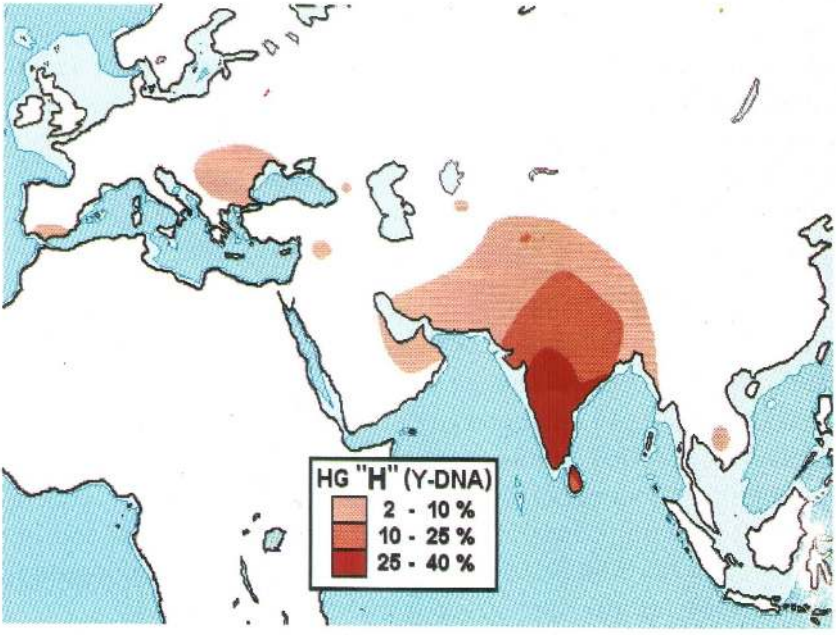
This genetic marker evolved from the first group of *Homo sapiens* to come out of Africa. One view is that the ancestors of this group migrated from Africa to the Middle-East, where about 45,000 years ago a new marker *M89* appeared among them. While many of the descendants of *M89* remained in the Middle-East, a group moved into India via Afghanistan, and another group moved into Central Asia through Iran (Wells, S. 2007: 209-10). The alternate view is that people with the genetic marker *M168* came out of Africa in a second major wave of migration, followed the African coastline through the southern Arabian Peninsula and arrived in India. From India one

branch went to the Middle East, another branch entered Indian inland, and the main branch proceeded towards Central Asia. (Sengupta, S. et al. 2006; Kivisild, T. 2003).

In India the distribution of this genetic marker has an overall frequency of 12.5%. It is present at a frequency of 18.1% among the Tribal groups and 9.6% among the castes (Cordaux, R. et al. 2004). It is present among the Panchamas and Sudras of Tamil Nadu and Andhra Pradesh (Zerjal, T. et al 2006; Sengupta, et al. 2006); and among the Koya (Dravidian) tribes at 27% (Kivisild, T. et al. 2003).

3. **Haplogroup H** (The Indian Marker I – Y Chromosome DNA marker *M69*): Evolutionary Line: *M168> M89> M69*

After the first Coastal Migration of the *M130*-clans through the peripheral regions of South Asia, we have already come across the migration of people with the genetic marker *M89*. A group among



Map showing the present-day distribution of Y-Chromosome Haplogroup H (Genetic Marker M69) in South Asia.
Picture Courtesy: Wikipedia.

the *M89* migrated to Iran and then into India. About 35,000 years ago a mutation identified as marker *M69* appeared on the male *M89* lineage which entered India. All their male descendants carried this *M69* genetic marker, which came to be known as an *Indian Marker*.

It is called the Indian marker because this mutation occurred in India, and it is concentrated there. It is the most frequently met Y-haplogroup among the tribal populations in South India; and represents the main Y-Chromosome Haplogroup of the indigenous Upper Palaeolithic inhabitants of India (Cordaux, R. et al. 2004; Sengupta, S. et al. 2006; Thanseem, I. et al. 2006). According to some geneticists this male ancestor may have been born in South Central Asia before their arrival in India (Wells, S. 2007: 211); but the fact that people carrying this marker is concentrated in India suggest that this mutation appeared in India.

The peoples of these clans migrated along the steppe highway from the Middle East, and then moved into India where they developed this mutation. Next to the African coastal migrants 55,000 years ago, these tribes are the major inland settlers of South Asia about 35,000 years ago. This group represents one of the main Palaeolithic populations of South Asia and were likely the first to establish significant settlements in India. Among the tribal populations of India, this is the most frequently found Genetic Marker (25-35%) (Cordaux, R et al. 2004: 231-5; Sengupta, S. et al. 2006).

Haplogroup H is common among the populations of India, Sri Lanka, Pakistan and Nepal. The highest frequencies of the Genetic marker M69 are found among 33% of the Dravidian speakers (Sahoo, S. 2006). In another study of the South Indian population, Genetic Marker M69 is present among 27.2 %, i.e. 110 among 405 people tested (Hammer et al. 2006). Among the Koya (a Dravidian language) speaking tribes in India *M69* is present in 71%; and in 10.3 % among the Sinhalese (Kivisild, T. *et al.* 2003). The Haplogroup H1, a subclade of this Haplogroup is present in India.

Outside India this Haplogroup is a major genetic group among the Balkan Romanis (gypsies) accounting for about 60% of the total Romani population. A variation of this Haplogroup is reported among

the Romani populations living in Bulgaria, Spain and Lithuania. *This supports the hypothesis of an Indian Origin of the gypsies*, when a group of their founder fathers migrated from India (Gresham et al. 2001; Moorjani, P. et al 2011).

4. **Haplogroup L** (The Indian Marker II – DNA marker *M20*): Evolutionary Line: *M168> M89> M9> M20*

About 50,000 years ago the Middle Palaeolithic humans with the Genetic Marker *M168* reached the Middle East from Africa. These people manifested a Y-Chromosome mutation called Marker *M89*. The carriers of this marker are identified as the *Middle Eastern clan.* About 40,000 years ago on the plains of the southern part of Central Asia or Iran another mutation identified as marker *M9* appeared on the *M89* lineage. The people carrying this marker *M9* are called the *Eurasian Clan*. The descendants of this clan were to expand their range to the ends of the Earth over the next 35,000 years.

Around 35,000 years ago, as the people with *M9* began to migrate eastward they encountered the great mountain ranges that define the south-central Asian highlands – the Hindu Kush running to the east, the Himalayas running from north-west to south-east and the Tien-Shan running south-west to north-east. The three ranges meet in the centre at the Pamir Knot in present day Tajikistan. Each of these ranges has peaks that soar to more than 5,000 metres or higher and as the world was in the grip of the Last Ice Age, crossing them would have been very arduous. It was because of these mountains that our Eurasian migrants had split into three groups – one moving to the north of the Hindu Kush, one group towards the east towards East Asia and the third to the south into India. How do we know this? The Y-chromosome distribution traces the route of these migrations (Spencer Wells 2003: 112-113).

The *M20* group which arrived in South India were said to be aggressive towards the Australoid people who were occupying that land at that time. Many geneticists who studied the maternal mt-DNA of the population believe that these *M20* eliminated the Australoid men and married their women. Many of the Australoid aborigines

withdrew into the jungles and mountains to avoid persecution, and led a secluded life.

Haplogroup L-M20 is currently present in the Indian population at an overall frequency of 7-15% (Basu *et al* 2003., Cordaux *et al.* 2004, Sengupta *et al.* 2006, Thamseem *et al.* 2006). It is especially frequent among the Dravidian upper and middle castes at about 17-19%; but somewhat rarer in Indo-Aryan upper and middle castes at about 5-6%. This suggest that it may have been (perhaps beside J2) the original Y-Haplogroup of the Indus Valley Civilization (Sengupta, S. *et al.* 2006).

Its highest frequency and diversity can be found in Baluchistan (West Pakistan), from where the agricultural creators of this civilization colonized the Indus Valley in the fourth millennium BCE (Qamar *et al.* 2002). Among the tribal groups in India, the presence of Haplogroup L is quite rare at about 5-6%, (Cordaux *et al.* 2004, Sengupta *et al.* 2006, Thamseem *et al.* 2006). This finding suggest that its arrival in South India was at a later stage, and it was not a Y-Haplogroup of the original Palaeolithic population of India.

Earlier studies (Wells, S. 2001, 2003) reported a very high frequency of Haplogroup L-M20 approaching 50%. This appears to have been due to extrapolation of data obtained from a sample of 84 Kallars, a Tamil speaking warrior caste from Tamil Nadu among whom 40 people showed the presence of the genetic marker *M20*.

5. **Haplogroup R2** (Central Asian group – DNA marker *M124*): Evolutionary Line: *M168> M89> M9> M207> M124*

We have already come across ancestors carrying the *M9* lineage moving in three directions: one into South Asia (*M20*), one into East Asia (*M175*), and the third one into Central Asia (*M45*). About 30,000 years ago an individual in the Central Asian region developed a new mutation *M207*. About 25,000 years ago one of the individuals among the M207 group who lived in South Central Asia developed another mutation marked by *M124*. The descendants of this group migrated southwards in a forked pattern. They settled in North India and from there moved to South by 25,000 years BP. *M124* and its ancestral line define the Haplogroup *R2* (Spencer Wells 2009: Genographic Project).

25,000 years ago this group formed part of the pre-Aryan population of South Asia. This marker is found in North India and Pakistan at frequencies of 5 to 10 percent (Spencer Wells 2007: 227). In South India the prevalence of this marker at present varies between 5-10 percent as well. A study by Dr. Kivisid and colleagues showed a high prevalence of *M124* among the Sinhalese of Sri Lanka (Kivisild, T. *et al.* 2003).

6. **Haplogroup J2** (Mesopotamian group – DNA marker *M172*): Evolutionary Line: *M168> M89> M304> M172*

This Marker *M172* associated with the Neolithic farming group of East Mesopotamia arrived in India from Elam/Southern Iran in Mesopotamia 10,000 – 9,000 years ago most likely via the Bolan Pass into Baluchistan and then into India. At present this marker is present among 32 % of the men in Khuzistan (ancient Elam), 28% of the Brahui (a Dravidian language) speaking men of Baluchistan, 19% of the men in Tamil Nadu and 15% of the men in Karnataka (Cavalli-Sforza 1994; Sengupta, S. 2006; Thiagarajah, Siva 2011: 122-123, 299). This genetic evidence supports the Proto-Elamo-Dravidian linguistic concept proposed by the Pennsylvania University academic David McAlpin (McAlpin, D. 1981).

7. **Haplogroup R1A1** (Indo-Aryan group – DNA marker *M17*): Evolutionary Line: *M168> M89> M9> M45> M207> M173> M17*

The genetic marker M17 is of East European origin which appeared 10,000 years ago in present-day Ukraine or southern Russia. His nomadic descendants would eventually carry this genetic marker from the steppes to places as far away as India and Iceland. This marker was brought to India by the Indo-Aryan people who migrated from Central Asia via Iran 3,800 years ago. M17 in India is found at a high frequency among the Punjabis, Gujaratis and Maharastrans speaking the Indo-Aryan languages. 72% of West Bengal Brahmins, the least contaminated of this group and 48% of Konkanastha Brahmins carry this marker. Among the Hindi speaking population of Delhi, this marker is present among 35% of the men. (Spencer Wells 2002: 167; Sengupta, S. *et al* 2006).

This group Arrived in North India from Central Asia via Iran 3,800 years ago and have spread far and wide all over South Asia. Major sections of them arrived in South India around 500 BCE, about 1,000 years after the onset of the South Indian Megalithic Culture. They now represent about 11% of the Tamil Nadu, 18% of the Kerala and 13% of the Sri Lanka male populations (Cavalli-Sforza, L.L. 1994; Kivisild, T. et al. 2003).

Apart from the above there are several other minor genetic groups among the Tamils. Among these minor groups: Haplogroup O3 (Chinese group – DNA marker *M122*) arrived in India from the Upper-middle Yellow River Basin in China in several waves through Tibet and Burma between 6,000 – 4,000 years ago. This group is believed to be responsible for the introduction of rice into India. They mainly occupy the sub-Himalayan regions – Nepal, Bhutan, Assam and Nagaland; and a high percentage is found in Bengal. Overall their presence is less than 1% in South India (Bing Su, Cavalli-Sforza *et al.* 2000).

Indian Castes and Tribal Populations

Genetic studies addressing the distinction between Indian castes and tribal communities have arrived at opposing conclusions (Petrraglia, M.D. & Allchin, B. 2007: 239): Those based on phylogenetic analysis have provided support for the common origin of Indian tribes and caste groups irrespective of their social hierarchy (Roychoudhury et al. 2000; Kivisild et al. 2003; Metspalu 2004), while those focussing on genetic distances have brought out the differences between tribal and caste groups, as a signal for the castes clustering with either Central Asian (Cordaux et al. 2004a) or with West Asian populations (Cordaux et al. 2004b). The reason for this conflicting evidence may lie in this simple explanation: The men who had been propagating castes arrived here at a later date, and because of the shortage of women who had accompanied them, had to marry the descendants of the native aboriginal women. However the descendants of the South Asian aboriginal women, 75% of whom had been occupying this land for more than 30,000 years, are distributed both among the castes and tribal populations. Some studies show that the contemporary

caste populations of Tamil Nadu are confounded by assimilation of subsequent immigrants in varying degrees of admixture before the caste system became too rigid (Kanthimathi, S. et al. 2008).

Conclusion

The important tribal and caste populations who speak the Tamil language in India have arrived in this land over a long period of time-span ranging from 55,000 YBP to 3,800 YBP. Some of them are the direct descendants of the earliest migrants from Africa who had arrived here on their long trek. Contrary to popular belief, there is no single Tamil race or a specific genetic or anthropological group which could be named as Tamils or even Dravidians. The Tamils are an assortment of peoples descended from different ages, from different places and from different genetic backgrounds who have come together to occupy this land and speak this distinguished language. It is also relevant to note that more than 70% of the present-day speakers of the Tamil language are the direct descendants of the ancestral populations who have occupied South Asia since 25,000 years ago. Their language may have evolved much later; but the people are of ancient stock.

(Because of space limitations I have restricted this study only to the most important groups among the male ancestral line. The female line offers another fascinating insight).

Bibliography

BASU, M.; SENGUPTA,S. *et al.* (2003): Ethnic India: A Genomic View, with Special Reference to Peopling and Structure, *Genome Research,* 2003, 13 (10).

BING SU, CHUNJIE XIAO, RANJAN DEKA, CAVALLI SFORZA et al. (2000): Y- Chromosome Haplotypes reveal Prehistoric Migrations to Himalayas, *Human Genetics,* Vol. 107, No.6, December 2000

CANN, R.L. *et al.* (1987): Mitochondrial DNA and Human Evolution, *Nature* 325: 31-36.

CAVALLI-SFORZA, L.L. *et al.* (1994): The History and Geography of Human Genes, Princeton University Press, Princeton, NJ.

CORDAUX, RICHARD; STONEKING, MARK *et al.* (2004a): Independent Origins of Indian Caste and Tribal Paternal Lineages, *Current Biology* 14 (3): 231-235.

CORDAUX, R.; DEEPA, E.; VISHWANATHAN, H.; STONEKING, M. (2004b): Genetic Evidence for the Demic diffusion of Agriculture to India, *Science* 304: 1125.

KANTHIMATHI, S.; VIJAYA, M. and RAMESH, A.(2008): Genetic Study of Dravidian Castes of Tamil Nadu, *Journal of Genetics*, Vol. 87, No.2: 175-179.

KIVISILD,T.; METSPALU, M. *et al.* (2003): The Genetic Heritage of the Earliest Settlers persists both in Indian Tribal and Caste Populations, *American Journal of Human Genetics*, 72:313-332.

McALPIN, DAVID (1981): Proto-Elamo-Dravidian: The Evidence and its Implications, American Philosophical Society, Philadelphia.

MELLERS, PAUL (2006): Going East: New Genetic and Archaeological Perspectives on the Modern Human Colonization of Eurasia, *Science* 313: 796 - 800.

METSPALU, M.; KIVISILD, T.; METSPALU, E., *et al.* (2004): Most of the extant mtDNA boundaries in South and South West Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans, *BMC Genetics* 5: 26 ff.

OPPENHEIMER, STEPHEN (2003): Out of Eden – The Peopling of the World, Constable & Robinson, London.

PETRAGLIA, M. D.(2003): The Lower Palaeolithic of the Arabian Peninsula: Occupations, Adaptations and Dispersals, *Journal of World Prehistory* 17: 141- 179.

PETRAGLIA, MICHAEL D. and ALLCHIN, BRIDGET (2007): The Evolution and History of Human populations in South Asia, Springer, The Netherlands.

QUAMAR, RAHEEL et al. (2002): Y-Chromosomal DNA Variation in Pakistan, *American Journal of Human Genetics*, Vol. 70(5): 1107-1124.

QUINTANA-MURCI, LUIS (1999): The mtDNA evidence for a Coastal Exodus from Africa, *Nature Genetics* 23: 437-441.

ROYCHOUDHURY, S., ROY, S., DEY, B. *et al.* (2000): Fundamental Genomic Unity of Ethnic India revealed by Analysis of Mitochondrial DNA, *Current Science* 79: 1182-1192.

SENGUPTA, S.; CAVALLI-SFORZA, LUCA, L. *et al.* (2006): Polarity and Temporality of High Resolution Y-Chromosome distribution in India identify both Indegenous and Exogenous Expansions and reveal Minor Genetic Influence of Central Asian Pastoralists, *American Journal of Human Genetics*, 78(2): 202-221.

STRINGER, C. and McKIE, R. (1996): The African Exodus, Jonathan Cape, London.

STRINGER, C. (2000): Coasting out of Africa, *Nature* 405: 24-27.

THAMSEEM, ISMAIL; THANGARAJ, KUMARASAMY et al. (2006): Genetic Affinities among the Lower Castes and Tribal Groups of India, *BMC Genetics,* 7: 42.

THIAGARAJAH, SIVA (2011): Peoples and Cultures of Early Sri Lanka, Tamil Information Centre, Kingston, UK.

VIGILANT, L. *et al.* (1991): African populations and the evolution of human mitochondrial DNA, *Science* 253: 1503-7.

WATSON, E. *et al.* (1997): Mitochondrial Footprints of Human Expansions in Africa, *American Journal of Human Genetics* 61: 691-704.

WELLS, R.S.; YULDASHEVA, N.; RUZIBAKIEV, R.; UNDERHILL, P.A. *et al.* (2001): The Eurasian Heartland: A Continental Perspective on Y-Chromosome Diversity, *Proceedings of the National Academy of Sciences* 98(18): 10244.

WELLS, SPENCER (2002, repr.2003): The Journey of Man, Princeton University Press, NJ.

WELLS, SPENCER (2007): Deep Ancestry, National Geographic, Washington DC.

ZERJAL, TATIANA; PANDYA, ARPITA; THANGARAJ, KUMARASAMY *et al.* (2006): Y-Chromosomal Insights into the Genetic Impact of the Caste System in India, *Human Genetics* 121 (1): 137-44.

2

The Origin and Spread of Dravidian Languages

Evidences in Favour of the Proto-Elamo Dravidian Hypothesis

(Paper Presented at the World Research Conference of Tamil Studies, SOAS, London 15 August 2013)

Abstract

To-date several theories about the origins and evolution of Dravidian languages have been put forward. Evidences from archaeology, migration of farming practices and population genetics support the hypothesis of a proto-Elamo-Dravidian origin and spread proposed by linguist David McAlpin. An outline of the available evidences is discussed in this article.

Introduction

Dravidians refer to the peoples of India, Pakistan, Sri Lanka and elsewhere who speak a language or a dialect belonging to the present-day South Indian Family of languages. The term Dravidian was first used by Rev. Robert Caldwell in 1856 as a generic name of the major language family next to Indo-Aryan spoken in the Indian subcontinent (Caldwell, R. 1856, repr. 1956: 3-6). But over the years the name Dravidian has been used to denote not only the languages, but also the broad ethnic groups of people who occupy South India; as well as the ancestral genetic groups from whom these people have descended.

This name was an adaptation of the Sanskrit term *Dravida* which was traditionally used to designate the Tamil people and Tamil language in some contexts, and in others to denote the South Indian

peoples. Caldwell cited several sources: *Manusmriti, Natyasastra,* and *Mahabharata* where the word *Dravida* is used to denote a people; Kumarilabhatta's *Tantravarttika* where *Dravida* is used as the name of a language; and *Dravidi* as a Prakrit language. He also stated that "by the adoption of this term 'Dravidian', the word 'Tamilian' has been left free to signify that which is distinctively Tamil" (Caldwell, R. 1956: 6).

The Prakrit forms of the Sanskrit Dravida are *Damila* which occurs in Pali literature, and *Demelu* in Ardhamagadhi; and the names *Damila, Demela, Dameda* and *Tamira* occurs in inscriptions and epigraphic records. They are all cognates of the Tamil name 'Tamizh', which refers to the Tamil language, Tamil country and Tamil people.

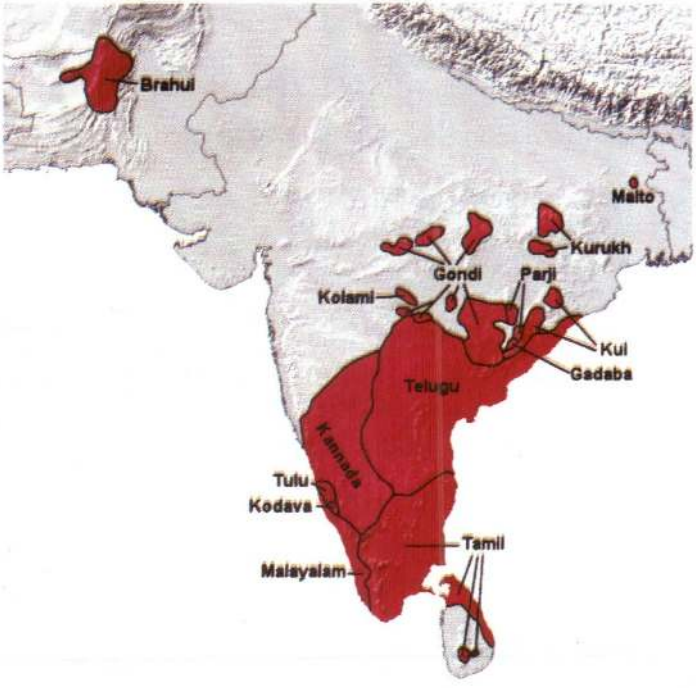
Languages belonging to the Dravidian family are spoken today mainly in South Asia, but migrant groups from this region have taken these languages to almost all the developed countries of the world. What is of interest to the historical linguist is the presence of some of the minor and less developed Dravidian languages in places outside South India; and the presence of related languages outside India.

The Dravidian languages of India are divided into four major groups: The Southern Group: 1. Tamil, 2. Malayalam, 3. Irula, 4. Kurumba, 5. Kodagu, 6.Toda, 7. Kota, 8. Badaga, 9. Kannada, 10. Koraga, 11. Tulu. The South-Central Group: 12. Telugu, 13. Gondi, 14. Konda, 15. Kui, 16. Kuvi, 17. Pengo, 18. Manda. The Central Group: 19.Kolami. 20.Naiki (Chanda), 21.Parji, 22.Naikri, 23. Gadaba-Ollari, 24. Gadaba-Salur, 25.Gadaba- Pottangi. The Northern Group: 26. Kurux, 27.Malto, 28. Koraga.

Apart from the 28 languages mentioned above and the 29th language Brahui spoken in Baluchistan, there are more than 50 minor Dravidian dialects scattered among the small communities all over India, Nepal, Pakistan and Iran, not developed enough to be classed as languages.

The four major Dravidian languages are concentrated in South India, each dominant in a separate linguistic state. Telugu has the largest number of speakers: an estimated 75 million is the language of Andhra Pradesh. Tamil, which has the oldest literature among the Dravidian languages spoken by about 65 million people is the language of Tamil Nadu. Kannada, the language of the state of Karnataka, has about 50 million speakers. Malayalam, the language of Kerala has an estimated 35 million speakers (Krishnamurthi,B.: 2003).

Of the minor languages, Gondi, Tulu and Kurukh are each spoken by more than a million speakers. Many of the others are spoken by tribal people. It is noteworthy that some of these languages are spoken by tribes in North India, in states like Madhya Pradesh, Orissa, Bihar and West Bengal where Indo-Aryan is currently the dominant language.



Present distribution of Dravidian Languages. Picture Courtesy: Wikipedia

The only Dravidian language spoken outside modern India is Brahui. It is spoken by the Brahui people living in Baluchistan, the south west region of Pakistan and the border regions of Afghanistan and Iran with Pakistan. The Pakistani cities Quetta, Khost, Kalat and Sibi which surround the ancient pre-historic site at Mehrgarh, are populated with Brahuis. At present there are 2.2 million people who speak the Brahui language. It has been significantly influenced and assimilated by the Iranian languages, especially Balochi. According to a recent UNESCO report, Brahui is one of the languages facing danger of extinction.

The *Mehrgarh Culture* site belonging to 9500-4600 years before present (YBP), is the earliest Neolithic archaeological site in South Asia, and the ancient people who still live in this region are Dravidians speaking the Brahui language. Mehrgarh was an early farming village, and its residents lived in mud brick houses, stored their grains in granaries, fashioned tools with local copper ore, and lined their large basket containers with bitumen. This site was discovered in 1974 by the French archaeologist Jean-Francois Jarrige. This culture is believed to be the fore-runner to the later Indus Valley Civilization and it has been suggested that Brahui might be a remnant of the language spoken in the Indus Valley.

The distribution of Dravidian languages in North India and Pakistan is of significance for tracing the history of the Dravidian family in the Indian subcontinent. Although in modern times the speakers of the major Dravidian languages are concentrated in peninsular India, as the Oxford Sanskrit professor Thomas Burrow and historical linguist Kamil Zvelebil has pointed out, it is a well-established and well supported hypothesis that Dravidian speakers were widespread throughout India, including the north-west region before the advent of the Indo-Aryans. (Burrow,T. 1960; Kamil Zvelebil, 2003: 698). The existence of a Dravidian India in the past is well validated through archaeology and anthropology, and a scientifically established fact through studies in population genetics (Allchin, F.R. 1995; Kennedy, K. 2000; Wells, S. 2002).

The Origin of Dravidian Languages

The date of origin of the Dravidian languages, as well as their subsequent development and their period of differentiation are not satisfactorily worked out as yet. This is mainly due to a lack of comparative linguistic research in the past into the Dravidian and related languages. Scholarly interest outside India has not been as widespread and sustained as in the case of the study of Indo-European languages, though it cannot be denied that a few devoted western scholars have been making important contributions to Dravidian historical linguistics over the last three decades. Three theories have been proposed as to the origin of the Dravidian languages.

Theory of Indigenous Origin

According to this theory the Dravidian languages are indigenous to India and evolved from ancient dialects spoken by the Ancestral Dravidians. Dr Bhadriraju Krishnamurti summarises this view as follows:

> Most of the proposals that the Proto-Dravidians entered the subcontinent from the outside are based on the notion that Brahui was the result of the first split of Proto-Dravidian and that the Indus Civilization was most likely to be Dravidian. There is no concrete evidence to credit Brahui with any archaic features of Proto-Dravidian. The most archaic features of Dravidian phonology and morphology are still found in the southern languages namely Early Tamil....For the time being, it is best to consider Dravidians to be the natives of the Indian subcontinent who were scattered throughout the country by the time the Aryans entered India around 1500 BCE (Krishnamurti, B. 2003: 5).

Theory of Language Migration

A popular theory in the past among many Dravidian linguists supported by a number of scholars including Robert Caldwell, Thomas Burrow and Kamil Zvelebil is that the Dravidian languages display similarities with the *Uralic* (Hungarian, Finnish) and the *Altaic* (Turkish, Mongol)

language groups, suggesting a prolonged period of contact among these languages in ancient times. Kamil Zvelabil pointed out that "a hypothesis has been gaining ground that posits a movement of Dravidian speakers from the northwest to the south and east of the Indian peninsula, a movement originating possibly from far away as Central Asia....the Uralic and the Altaic connection is definitely most probable and most promising" (Burrow, T. 1944; Zvelabil, K. 2003:698). On another occasion Zvelabil himself had stated that "All this is in the nature of speculation; a truly convincing hypothesis has not even been formulated yet" (Zvelabil, K. 1990:123). However this theory of an Ural-Altaic connection has now been rejected as impossible by several specialists in Uralic languages (Zvelabil, K. 1970, Krishnamurti, B. 2003: 43).

Sumerian, Elamite and Akkadian were ancient languages spoken in ancient Mesopotamia. During the 1960s Professor Sathasivam (Colombo and Berkley) discovered a strong relationship between the extinct Sumerian language and the extant Dravidian languages and concluded that Sumerian belonged to the same family of Dravidian languages (Sathasivam, A. 1965; 1966). Dr. Loganathan from the University of Malaysia, has demonstrated the lexical correspondences between Sumerian and Dravidian as well as Akkadian and Dravidian (Loganathan, K. 1989), reiterating the ancient linguistic connection between Mesopotamia and Dravidian India. The affinity between Elamite and Dravidian languages identified by the Pennsylvania University academic Dr. David McAlpin, (McAlpin 1981) which is supported by additional archaeological and genetic evidences, is discussed below.

Theory of Mixed Origin

It is very likely that the proto-Dravidian is a composite language of mixed origins. According to studies in population genetics the ancestors of about 70% of the speakers of the present-day Dravidian languages have been living in South Asia for the past 30,000 years. We do not know the language or languages they spoke during the Upper Palaeolithic and Mesolithic periods. When the Neolithic cultivators from the Fertile Crescent brought a language along with the passage

of agriculture 9,000 years ago and when the native hunter-gatherers took up to agriculture, the indigenous language of the natives would have mixed up with the new agricultural language, contributing to the origins of this composite proto-Dravidian language (Thiagarajah, Siva 2011: 122-23).

The Proto-Elamo-Dravidian Hypothesis

According to this hypothesis the Elamo-Dravidian languages belonged to an ancestral language family which includes the living Dravidian languages of India and Pakistan, in addition to the extinct Elamite languages spoken in ancient Elam which is now south-western Iran, as well as in ancient Persia. The common ancestral language of this family is named Proto-Elamo-Dravidian.

Today, Dravidian languages are essentially restricted to South India and Sri Lanka. The Proto-Elamo-Dravidian hypothesis proposes that they originated in the Iranian province of Elam and were once spoken over a much larger area including Iran, Pakistan, Afghanistan and the whole of India (McAlpin, D. 1974, 1981).

The Linguistic Evidence

In 1853, Edwin Norris writing about the Elamite inscription at Behistun made numerous comparisons with Tamil, a Dravidian Language (Norris, E. 1855: 6-7). In 1856, when Caldwell published his *Comparative Grammar of Dravidian or South Indian Family of Languages,* he incorporated Norris's work, placing it as the first suggested affiliation for Dravidian and discussing the connection at some length (Caldwell,R. 1856: 43-45; repr.1974: 65-67).

Over the years several scholars of Elamite studies have noted this linguistic connection with the Dravidian. Notable among them were Georg Husing (1901, 1910); Vilhelm Thomsen (1910); Alfredo Trombetti (1913); Ferdinand Bork (1925) and Pedersen (1962:128). Until 1974, the best description of the relationship between Elamite and Dravidian was given by the Russian scholar I.M. Diakonoff (1967). After discussing the Elamite language at some length, he

presented in six pages all the arguments for a genetic relationship between the two languages, along with a list of seven possible etyma (Diakonoff, I.M. 1967: 85-112).

It was left to the unstinting efforts of David W. McAlpin, then Associate Professor of Dravidian Languages and Linguistics at the University of Pennsylvania, to provide the evidence for this historic connection. He coined the term Elamo-Dravidian to indicate this connection and gave the name 'Proto-Elamo-Dravidian' to the ancestral common mother tongue. He presented his preliminary paper "Toward Proto-Elamo-Dravidian" which appeared in the journal *Language* in 1974. This was followed by "Elamite and Dravidian, Further Evidence of Relationships" which appeared in *Current Anthropology* in 1975. His definitive work, "Proto-Elamo-Dravidian" was published in 1981 in Philadelphia (McAlpin, David 1981).

McAlpin identified several similarities between Elamite and Dravidian. According to McAlpin 32% of the Dravidian and Elamite vocabulary are cognates and possible cognates. Elamite and Dravidian possess similar second person pronouns and parallel case endings. They have identical derivatives, abstract nouns and the same *verb stem+tense marker+personal ending* structure. Based on this he postulated a Proto-Elamo-Dravidian language from which the Elamite and Dravidian languages had evolved (Mc Alpin, David 1981).

Kamil Zvelbil who had reviewed McAlpin's study in greater depth had commented that more additional work need to be done in this field to shift it from the realm of hypothesis to establish it as a fact acceptable to all (Zvelebil, K.,1990).

Apart from this linguistic affinity, the Elamo-Dravidian concept is validated on evidence provided by two other disciplines: (1). Archaeological studies showing that agriculture developed in the Near East spreading to Mehrgarh, the Indus Valley region and then into India, while indicating that the Elamo-Dravidian languages too had spread along with the passage of agriculture.

(2). Genetic studies of recent years revealing the migratory paths of the people associated with agriculture from Elam to India (Cavalli Sforza, L.L. 1995: 177; Kivisild, T. et al. 2002: 215-222; Sengupta, S. et al: 2006).

The Archaeological Evidence

The foundations of modern civilization begin with the development of agriculture 11,500 years ago when the hunter-gatherer groups of the past settled along the river banks and began a transformation from a hunting-gathering economy to a food producing economy (Barker, G. 2009). The Last Ice Age have come to an end, a warmer climate have set in, rain was plentiful and the rivers were overflowing This stage marked the onset of the Neolithic or New Stone Age, which is traditionally considered as the last part of the Stone Age (Bellwood, P. 2004). Agricultural activities began independently in the Middle East and in China.

Evidence points to the Fertile Crescent of the Middle East, as one of the earliest sites of planned sowing and harvesting of plants that had previously been gathered in the wild. Elam, both Susania and the Zygros regions form the south-eastern end of the Fertile Crescent and took part in this accomplishment. After several centuries of steady development, by 9,000 BCE this revolution in food production became an accomplished fact in this region (Harris, D.R. & Gosden, C. 1996; Barker, G. 2009).

The result was the Neolithic farming village on this West Asian model, cultivating primarily wheat and barley with some lentils and raising sheep, goats and cattle. This *wheat-barley-sheep-goat-cattle complex* of agriculture and pastoralism spread all over South Asia, and became the dominant agricultural pattern in large parts of the Old World.

The spread of agriculture has been associated with the dispersal of a proto-Dravidian language (Cavalli-Sforza 1988; Bellwood & Renfrew 2002; Fuller, D.Q. 2007: 393-443). The Elamo-Dravidian languages are believed to have originated in the Elam province- that

is around Zagros Mountains and Southwest Iran (Ruhlen 1991). Along with the spread of agriculture the Proto-Elamo-Dravidian language too, spoken by the Elamites in Southwestern Iran, spread eastward with the movement of these farmers from this region to the Indus Valley and the Indian subcontinent (Ruhlen, M. 1991; Cavalli-Sforza et al. 1994, 1996, 2000; Renfrew C. 1996).

The new farming villages and towns of Persia (Iran and Afghanistan) carried with them Elamo-Dravidian names. This language was spoken in the region for thousands of years until Elamite became the official language of the Persian Empire, even as late as, from the Sixth to the Fourth centuries BCE. Early twentieth century scholars who studied the place names in Persia, mistook these Elamo-Dravidian names for early Dravidian, and took this as evidence for the spread of Dravidian into Persia from India in ancient times (Ramaswami Aiyar, L.V.: 1929-30).

The agricultural revolution soon spread across the Persian border into Ancient India (Pakistan+India). Wheat, barley and jujube were domesticated in the Indian Subcontinent by 8,000 BCE, soon followed by domestication of cattle, sheep and goat in the Mehrgarh Culture by 8,000 BCE (Harris et al. 1996; Barber, Z. 1996). This period also saw the first domestication of the elephant (Gupta, A.K.: 2004). Agro-pastoralism in India included threshing, planting crops in a row and storing grain in granaries. Archaeological evidence indicates that rice was cultivated in the Gangetic plains and was part of the Indian diet by 5,000 BCE (Nene, Y.L.: 2005; 85-106). By the third millennium BCE cotton was also cultivated, and agricultural communities became widespread in areas like Kashmir and Karnataka regions (Harris, D.R. *et al.* 1996; Nene, Y.L. 2005). By 2,000 BCE rice cultivation had arrived in South India and very likely in Sri Lanka as well. How large (or small) the population movements were that had accompanied the advance of agriculture could only be answered through a genetic survey.

Genetic as well as archaeological evidence indicate that the 'Elamo-Dravidian' agriculturalists were responsible for bringing farming from West Asia to Mehrgarh and the Indus Valley. Later

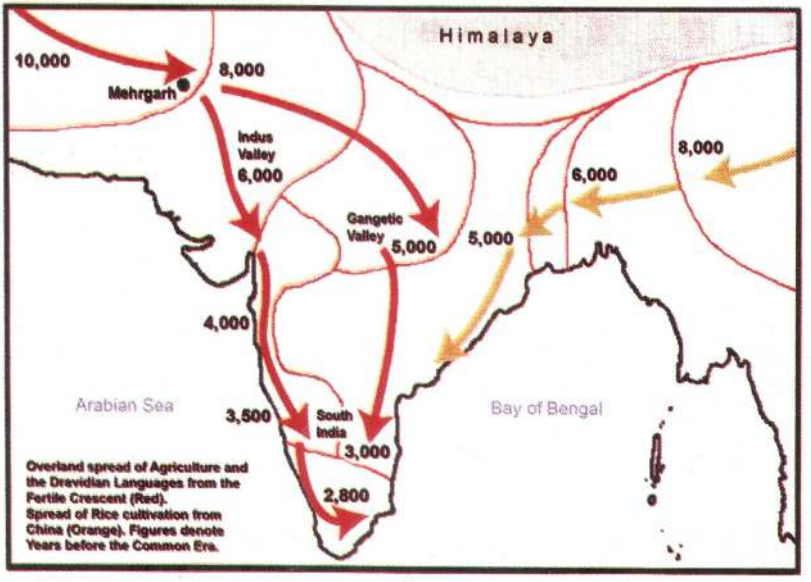
evidence of extensive trade between Elam and the Indus Valley Civilization suggests ongoing links between the two regions. Similarities have been noted between the Harappan script which has not been conclusively deciphered, and early Elamite script (Cavalli-Sforza 1994, 2000; Ruhlen, M. 1991; Quintana-Murci, L. *et al* 2001).

The Pre-Indus Mehrgarh Culture and The Spread of Agriculture

Situated near the Bolan Pass, to the west of the Indus River valley, between the present-day Pakistani cities of Quetta, Kalat and Sibi, Mehrgarh was discovered in 1974 by the French archaeologist Jean-Francois Jarrige. Excavated continuously between 1974 and 1986, Mehrgarh became one of the most important Neolithic (8000 BCE to 2600 BCE) sites in archaeology.

The earliest settlement at Mehrgarh, in the northeast corner of the 495-acre site, was a small farming village dated between 7,500 BCE and 5,500 BCE (Jarrige, J-F.:1982; Hirst, K.K.:2005). Early Mehrgarh residents lived in mud brick houses, stored their grain in granaries, fashioned tools with local copper ore, and lined their basket containers with bitumen. They cultivated barley, emmer wheat, einkorn, jujubes and dates. They herded sheep, goats and cattle. Residents of the later period 5.500 BCE to 2,600 BCE) put much effort into crafts, including flint knapping, tanning, bead production, and metal working. The site was occupied continuously until about 2,600 BCE (Possehl, G.L.: 1996).

Mehrgarh is now seen as a precursor to the Indus Valley Civilization. As professor Ahmed Hasan Dani of Islambad Quaid-e-Azam University points out "There we have the whole sequence, right from the beginning of settled village life progressing to the urban". The people of Mehrgarh also had contacts with contemporaneous cultures of North Afghanistan, North-Eastern Iran and South-Central Asia (Kenoyer, J.M.: 2005). Although Mehrgarh is unoccupied at present, the Brahui speaking Dravidian people continue to live in the surrounding cities of Quetta, Kalat and Sibi.



Map showing the passage of Agriculture via Mehrgarh in 8,000 BCE and the arrival of Rice Cultivation from China in 5,000 BCE.

The Language Associated with the Spread of South Asian Agriculture

The imprints that the Dravidian languages had left behind in the field of agriculture and irrigation could be found stamped all over the Indian subcontinent. The domesticated crops, the implements used in agriculture, the art of irrigation from rivers and lakes, the processing and storage of farm produce – all have a substratum of Dravidian names attached to them (Southworth, F.C. 2005, 2006: 121-150).

"Dravidian languages have provided the Indo-Aryan languages of North India with many words for plants, animals and objects native to the Indian scene" (Encyclopaedia Britannica: 15th Edition, 1984).

Several ancient agricultural crops which were first domesticated in India carried with them a Dravidian name. For example, Asian rice (*Oryza Sativa)* when it was first domesticated and cultivated in the Gangetic plains during 5000 BCE by the Dravidian speakers who

were occupying that region at that time, was given a Dravidian name *arici /aris*. Rice is derived from the Old Dravidian *arici*, a language closer to Tamil; it is *arriz / arruz* in Arabic; *orez* in Hebrew; *oryza* in Greek; *oriza* in Latin; *ris* in Old French; *riz* in New French; *riso* in Italian; *rice* in English; *arroz* in Portuguese; *arroz* in Spanish; and *arroz* in Brazilian. They also gave cotton to the world. This is their lasting legacy.

The migration of people from the regions of Zagros and Khuzistan in South-West Iran (ancient Elam) into Ancient India discussed below is well supported by evidence furnished by modern genetic and archaeological studies. As summarised by the world renowned population geneticist Luca Cavalli-Sforza: *"It is likely that the early Dravidian languages were spoken from the western borders of Iran into and throughout the whole of India, and were first introduced by Neolithic cultivators from nine thousand years ago"* (Cavalli-Sforza, L.L. 1995: p.177).

The Brahui Controversy

The Brahui language spoken in Baluchistan is generally considered to be a remnant of the proto-Elamo-Dravidian left behind on its passage from Iran to India. However there are some objections to this view on linguistic grounds.

It has been argued on linguistic grounds that the absence of any Avestan (Old Iranian) loan words in Brahui suggest that Brahui migrated to Baluchistan at a later date. The main Iranian contributor to Brahui vocabulary is Balochi, a western Iranian language which arrived in this region from the west around 1000 CE (Elfenbein, J. 1987). Brahui language shares similar vocabulary with Kurukh and Malto spoken in Central India. This suggests the possibility that Brahui was originally spoken near Central India (Krishnamurti, B. 2003: 27, 142). Nevertheless this argument is also applicable to the converse: that Kurukh and Malto migrated from the West to Central India. The presence of Baluchi words in modern day Brahui does not signify anything; they had a thousand years of contact for language transferance.

Furthermore there is no historical evidence suggesting the relocation of Brahui speaking people from Central India to Baluchistan as recent as 1000 CE. Historical records indicate that the Brahuis are descendants of the Turko-Iranian tribes from West Asia (Hughes-Buller 1991).

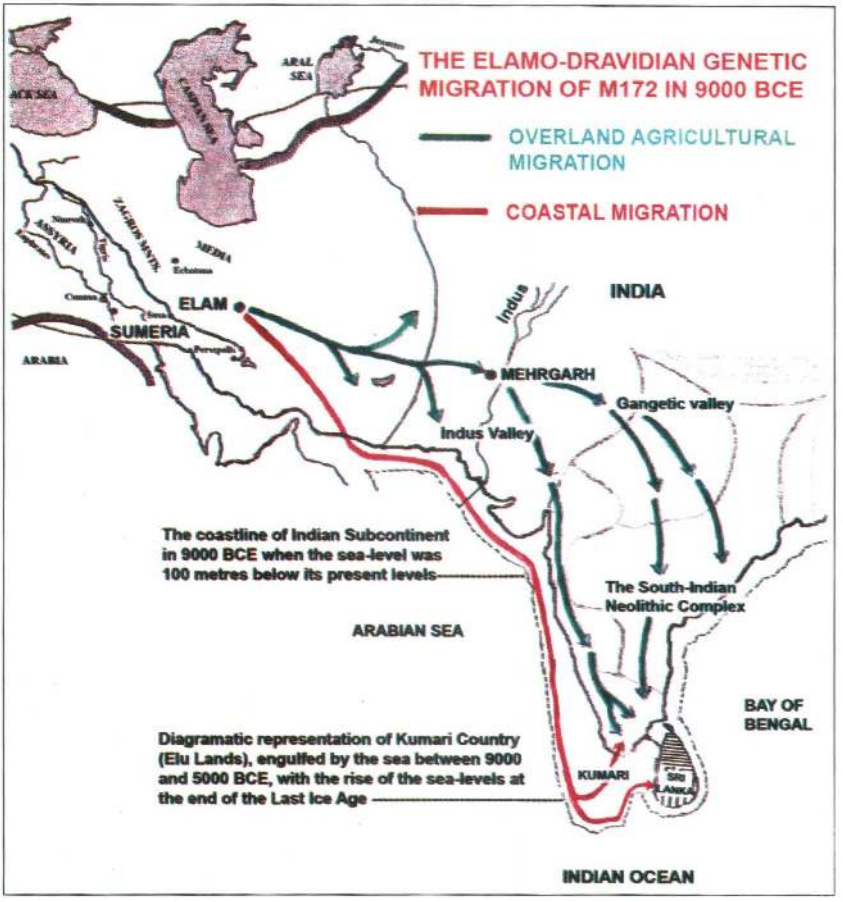
The genetic picture indicates that a good proportion of the males who speak the Brahui language are more related to the Elamites than to the North and Central Indians. The genetic marker *M172* present among 32% of the men in Khuzistan (Old Elam) is present among 28% of the Brahui speaking men of Baluchistan. In North and Central India the frequency of this marker among the men is around 11%. This suggests the migratory path of this genetic marker is from West to East – from Iran towards India. This is discussed below.

The Genetic Evidence

Overland And Coastal Migrations Of Genetic Marker M172

It appears that the first wave of Proto-Elamo-Dravidians moved from Elam (Mesopotamia) all the way to southernmost India as soon as the Ice Age came to an end 12,000 years ago, long before the movement of the agriculturists. As the Last Ice Age was coming to an end in the equatorial regions around 10,000 BCE the Elamites would have commenced on this journey. From Elam they would have moved south to the shores of the Arabian Sea and then migrated along the sea-beach towards South India. To the pre-historic population the sea-beach was their natural highway. They would have arrived in the Kumari Land of southernmost India around 9,000 BCE (even if we allow a thousand years for this trek), a land later to be consumed by the sea around 5,000 BCE. However, some of the people and the language they had brought with them would have penetrated into the interior.

The Proto-Elamo-Dravidian language on its overland route brought by the agriculturists would have accumulated several words from the Austro-Asiatic Munda languages in its earlier passages; and also would have exchanged a large amount of Prakrit and Sanskrit words during its later stages, changing its character altogether. Thus,



The Coastal Migration Hypothesis (Siva Thiagarajah 2011) is supported by two areas of evidence: 1. Linguistic Evidence: Among all the Dravidian languages the Sangam Tamil is closest to Middle Elamite. 2. Genetic Evidence: Among all the present-day people of India, the 'Elamo-Dravidian' genetic marker M172 is higher among the Tamil and Kerala populations.

one would expect the Tamil language, the last to arrive by the overland route to be the most adulterated of them all. However, Classical Tamil Language, the southernmost of all Dravidian languages, is the least adulterated; and *among all the Dravidian languages, the Sangam Tamil is most close to Middle Elamite than any of its peers* (David Mc Alpin 1981). Even Brahui, which is a neighbour to Elamite is heavily adulterated with Balochi and is further away from Elamite on its linguistic affinity. Thus, the language most close to

the Proto-Elamo-Dravidian among the modern Dravidian languages is Tamil. This could be possible only if the language had arrived by an alternative route via the sea or along the coast.

Throughout history, people of the southernmost Tamil Nadu - the Pandyan Country, and the Tamils of northern Sri Lanka are reputed to have spoken a dialect similar to the *Sangam Tamil*, which was in vogue 2,500 years ago. While this has now died out in Tamil Nadu, the people of Jaffna in northern Sri Lanka still continue to speak this dialect.

The most conclusive evidence for a coastal migration comes from the genetic picture of the present day Tamil and Kerala populations. The presence of the Fertile Crescent (Proto-Elamo-Dravidian) male genetic marker M172 is present among 32 % of the men of Khuzistan in southern Iran (ancient Elam). This marker is present among 28% of the Brahui speaking men of Baluchistan and among 11% of the men in North India. Through an overland route alone one would expect a gradual fall in the distribution of this genetic marker as it migrated to the southernmost regions of the peninsula. However, the frequency of this genetic marker M172 reaches 19 and 21% among the male populations of Tamil Nadu and Kerala, and 15% among those in Karnataka. (Quamar, R. 2002; Cavalli-Sforza 1994, 2004; Sengupta, S. 2006; Thiagarajah, S. 2011: 122). This indicates that there was an alternate passage as well for the arrival of the Proto-Elamo-Dravidians in the extreme south.

From about 8,000 BCE until 4.000 BCE, there had been a steady movement of the Proto-Dravidian languages from the northwest to the southeast of India (Cavalli-Sforza 1995). *By 3,000 BCE four distinct dialect groups: North Dravidian, Central Dravidian, South Central Dravidian and South Dravidian have evolved. It is from the South Dravidian, that Tamil, Kannada and Malayalam languages have descended* (Krishnamurti, B. 2003).

The Last Refuge of the Dravidian Languages

The major ancestors of the Dravidians have occupied South Asia for more than 30,000 years. For almost seven thousand years from 10,000 BP until 4,000 BP people speaking the proto-Dravidian languages

extensively occupied the lands of Iran, Afghanistan, Pakistan, India and Sri Lanka. The arrival of pastoral nomads from the Central Asian Steppes to the Iranian Plateau, 4,000 years ago, brought with it the Indo-Iranian branch of the Indo-European language family, which eventually replaced Dravidian languages in Iran and Afghanistan; and most of Pakistan and northern India, perhaps by an elite-dominance process (Renfrew, C. 1996; Cavalli-Sforza et al. 1994, Cavalli-Sforza 1996, 2000; Ruhlen, M. 1991). The Dravidian speaking peoples, who once dominated the entire South Asian landscape, became confined to South India, north-eastern Sri Lanka and to several isolated villages and hamlets in Pakistan and North India (Thiagarajah, Siva 2011: 121).

Bibliography

ALLCHIN, F.R. (1995): The Archaeology of Early Historic South Asia; Cambridge University Press, Cambridge.

BARBER, ZAHEER (1996): The Science of Empire: Civilization and Colonial Rule in India; State University of New York Press.

BARKER, GRAEME (2009): The Agricultural revolution in Prehistory: Why Foragers become Farmers, Oxford University Press.

BELLWOOD, P. & RENFREW, C. (2002): Examining the Farming / Language Dispersal Hypothesis, McDonald Institute, Cambridge.

BELLWOOD, PETER (2004): The First Farmers: The Origins of Agricultural Societies; Wiley- Blackwell, London.

BRAIDWOOD, ROBERT (1960): The Agricultural Revolution; *Scientific American* 203, Sept. 1960.

BURROW, THOMAS (1960): Sanskrit and the Pre-Aryan Tribes and Languages, *Indo-Asian Culture*, April 1960.

CALDWELL, ROBERT (1856; repr.1956): A Comparative Grammar of the Dravidian or South-Indian Family of Languages; University of Madras.

CAVALLI-SFORZA, L.L. et al. (1988): Reconstruction of Human Evolution: Bringing together Genetic, Archaeological and Linguistic Data; PANS 85(16).

CAVALLI-SFORZA LL., PIAZZA A., MENOZZI P. (1994): The History and Geography of Human Genes, Princeton Univ. Press, Princeton, NJ.

CAVALLI-SFORZA, LUIGI LUCA (1995): The Great Human Diaspora: The History of Diversity and Evolution; Addison-Wesley Pubishing, USA.

CAVALLI-SFORZA, L.L. (1996): The spread of Agriculture and Normadic Pastoralism: Insights from the Genetics, Linguistics and Archaeology. In HARRIS DR (ed): The Origins and Spread of Agriculture and pastoralism in Eurasia, Smithsonian Institution, Wasington DC.

CAVALLI-SFORZA (2002): Genes, Peoples and Languages, Penguin, London.

DIAKONOFF, I.M. (1967): Elamskij Jazyk, *Jazyki dernej perednej Azii,* (in Russian), Moscow.

ELFENBIEN, JOSEF (1987): A Periplus of the Brahui Problem, *Studia Iranica* 16(2): 215-232.

FULLER, DORIAN Q. (2007): 'Non-human genetics, Agricultural origins and Historical linguistics in South Asia' in PETRALGIA,M. & ALLCHIN,B. (eds.): The Evolution and History of Human Populations in South Asia; Springer, Netherlands

GUPTA, ANIL K. (2004): Origin of Agriculture and Domestication of Plants and Animals linked to Early Holocene Climate Amelioration; *Current Science*, Vol.87, No.1, July 2004.

HARRIS, DAVID R. & GOSDEN, C. (1996): The Origins and Spread of Agriculture and Pastoralism in Eurasia; Routledge, London.

HIRST, KRIS K. (2005): Mehrgarh: Guide to Archaeology

HUGHES-BULLER, R. (1991): Imperial Gazetteer of India: Provincial Series: Baluchistan, Sang-e-Meel, Lahore.

JARRIGE, J-F.(1982): Excavations at Mehrgarh. In POSSEHL, G.L.(ed.): Harappan Civilization; Oxford and IBH Publishing Co. New Delhi.

KENNEDY, K.A.R. (2000): God-Apes and Fossil Men: Palaeoanthropology of South Asia, University of Michigan Press, USA.

KENOYER,J; MARK & KIMBERLEY HEUSTON (2005): Ancient South Asian World; Oxford Univ. Press, Oxford.

KIVISILD, TOOMAS *et al.* (2002): The Genetics of Language and Farming Spread in India *in* BELLWOOD, P & RENFREW, C. (eds.) Examining the Farming/Language Dispersal Hypothesis, McDonald Institute, Cambridge.

KORISETTAR, R., VENKATASUBBAIAH, P.C., FULLER, D.Q.(2001):'Brahmagiri and Beyond: The Archaeology of the Southern Neolithic', in Settar,S.& Korisettar, R. (eds.):Archaeology of South Asia; Manohar, Delhi.

KRISHNAMURTHI, BADRIRAJU (2003): Dravidian Languages; Cambridge University Press.

LOGANATHAN, K. (1989): Lexical Correspondences between Sumerian and Dravidian, *Journal of the Institute of Asian Studies*, March 1989, Chennai.

McALPIN, D.W. (1974): Toward Proto-Elamo-Dravidian Language 50: 89-101.

McALPIN, DAVID W. (1981): Proto-Elamo-Dravidian: The Evidence and its Implications; American Philosophical Society, Philadelphia

NENE,Y.L. (2005): Rice Research in South Asia through the Ages: *Asian Agri- History*, Vol.9, No.2, 2005: 85-106.

POSSEHL, GREGORY L. (1996): Mehrgarh *in* Oxford Companion to Archaeology, (ed.) BRIAN FAGAN, Oxford University Press, Oxford.

QUAMAR, RAHEEL *et al.* (2002): Y-Chromosomal DNA Variation in Pakistan, *American Journal of Human Genetics*, Vol. 70(5).

QUINTANA-MURCI, L. *et al.* (2001): Y-Chromosome Lineages trace Diffusionof People and Languages in South-western Asia, *American Journal of Human Genetics,* Vol. 68(2): 537-542

RENFREW, COLIN (1987): Archaeology and Language: The Puzzle of Indo European Origins; Jonathan Cape, London.

RENFREW, COLIN (1996): *Languages Families and the Spread of Farming* in HARRIS, D.R. : The Origins and Spread of Agriculture and Pastoralism in Eurasia, Smithsonian Institution Press, Washington DC.

RENFREW, C. & BAHN, P. (1996): Archaeology: Theories, Methods and Practice; Thames & Hudson, London.

REDMAN, CHARLES L. (1978): The Rise of Civilization: From Early Farmers to Urban Society in the Ancient Near East; San Francisco.

RUHLEN,M. (1991): A Guide to the Worlds' Languages; Stanford University Press, Stanford.

RAMASWAMY AIYAR (1929-30): Dravidic Place Names in the Plateau of Persia; *Quarterly Journal of the Mythic Society*; Vol. XX, Bangalore.

SANKALIA, H.D. (1974): Prehistory and Protohistory of India and Pakistan; Postgraduate and Research Institute, Deccan College, Poona.

SATHASIVAM, A. (1965): Sumerian – A Dravidian Language, Berkley, USA.

SATHASIVAM, A. (1966): The Dravidian Origin of Sumerian Writing: *Proceedings of the First International Conference Seminar of Tamil Studies,* Vol. II, Madras, April 1966.

SCHMANDT-BASSART (1978): An Archaic Recording System and the Origin of Writing; Syro-Mesopotamian Studies Vol.1, No.2:1-32.

SENGUPTA, SANGAMITRA (2006): Polarity and Temporality of High Resolution Y-Chromosome Distributions in India: *American Journal of Human Genetics*, 78(2006): 202-221.

SOUTHWORTH, F.C. (2005): Linguistic Archaeology of South Asia, Routledge, London.

SOUTHWORTH, F.C. (2006): Proto-Dravidian Agriculture in OSADA,T. ed. *Proceedings of the Harvard-Kyoto RIHN Symposium,* Kyoto.

THIAGARAJAH, SIVA (2011): Peoples and Cultures of Early Sri Lanka, Tamil Information Centre, Kingston, UK.

WELLS, SPENCER (2002, repr. 2003): The Journey of Man; Random House, New York.

WELLS, SPENCER (2007): Deep Ancestry: Genographic Project, National Geographic, Washington D.C.

ZVELEBIL, KAMIL (2003): 'Dravidian Languages', *The New Encyclopaedia Britannica,* 22.

ZVELEBIL, KAMIL (1970): Comparative Dravidian Phonology, Mouton, The Hague. (Page 22 of this book contains a list of articles which oppose the Uralic - connection theory).

3

The Evolution of Tamil and other Dravidian Languages

Introduction

The term Dravidians refer to the peoples of India, Pakistan Sri Lanka and elsewhere who speak a dialect belonging to the present-day South Indian Family of languages. The name Dravidians is also used to denote not only the ethnic groups of people who occupy South India and speak the Dravidian languages, but also to the ancestral genetic groups from whom these people have descended.

As early as the second decade of the 19th century, the existence of the Dravidian language family was suggested by Alexander D. Campbell in his work *Grammar of the Teloogoo Language* from the close relationship between the two languages and indicated they have descended from a common ancestor (Campbell, A.D. 1849). William Carey, the Baptist Missionary linguist of Serampore in Bengal, felt that the languages of South India – Tamil, Telugu, Malayalam, Tulu, Kannada etc. formed a different groups from those of North India. But it was Friedrich Max Muller, the great Sanskrit scholar and the first editor of the *Rig Veda* who, in the fifties of the 19th century first separated the South Indian languages as being members of a distinct non-Indo-European ancestral family (Chatterji, S.K. 1965: 7). It was the publication of Robert Caldwell's *Comparative Grammar of the Dravidian or South-Indian Family of Languages* in 1856 considerably expanded the Dravidian domain establishing it as one of the major language groups of the world.(Caldwell, R.: 1856).

The word *Dravidian* is an Anglo-Indian expression; a word created by British scholars in India, and is a hybrid consisting of the Sanskrit word *Dravida* and the English adjectival suffix *–ian*,

which itself is from Latin. The term Dravidian was first used by Rev. Robert Caldwell in 1856 as a generic name of the major language family next to Indo-Aryan spoken in the Indian subcontinent (Caldwell,R.: 1856, repr. 1956: 3-6). This name was an adaptation of the Sanskrit term *Dravida* which was traditionally used to designate the Tamil language and Tamil people in some contexts, and in others to denote the South Indian peoples. Caldwell cited several sources: *Manusmriti, Natyasastra,* and *Mahabharata* where the word *Dravida* is used to denote a people; Kumarilabhatta's *Tantravarttika* where *Dravida* is used as the name of a language; and *Dravidi* as a Prakrit language. He also stated that "by the adoption of this term 'Dravidian', the word 'Tamilian' has been left free to signify that which is distinctively Tamil" (Caldwell, R.: 1856: 6).

The Prakrit forms of the Sanskrit Dravida are *Damila* which occurs in Pali literature, and *Demelu* in Ardhamagadhi and the names *Damila, Dhamila, Demela, Dameda* and *Tramira* occur in inscriptions and epigraphic records. They are all cognates of the Tamil name 'Thamizh', which refers to the Tamil language, Tamil country and Tamil people. Epigraphic evidence of an ethnic group termed as such is found in ancient Sri Lanka, where a number of inscriptions datable from fifth century BCE through radiometric dating have come to light.

It is a well established and well supported hypothesis that the Dravidian speakers must have been wide spread throughout India, including the northwest region during the second millennium BCE. This is clearly established because of a number of features of the Dravidian languages appearing in the Rigveda, the earliest known Indo-Aryan literary work indicating that the Dravidian Languages must have been present in the areas during the invasion of the Indo-Aryans from the north. Several scholars have demonstrated the influence of Dravidian on the Indo-Aryan languages in phonology, syntax and vocabulary.

Languages belonging to the Dravidian family are spoken today mainly in South Asia, but migrant groups from this region have taken these languages to almost all the developed countries of the

world. What is of interest to the historical linguists is the presence of some of the minor and less developed Dravidian languages in places outside South India and the presence of related languages outside India.

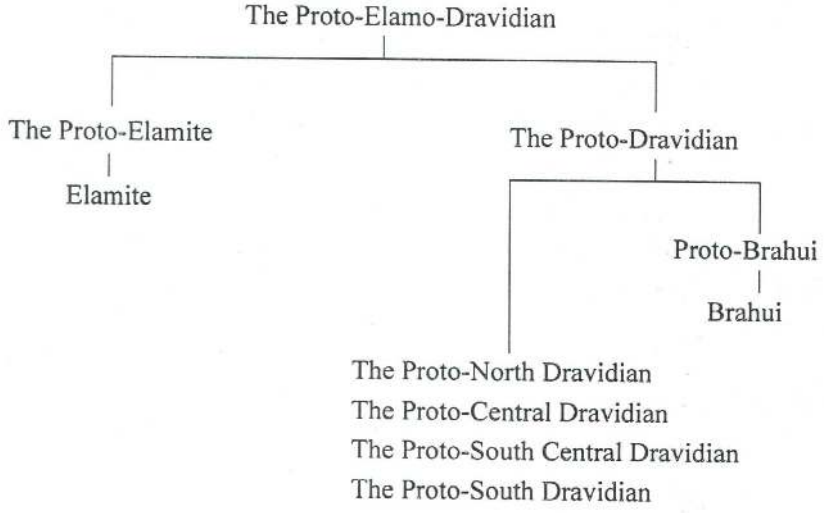
Regional Development of Dravidian Languages

All the languages spoken in the world today are borne from its parent language, which in turn had its own parent. It takes thousands of years for these daughter languages to evolve from a parent and the process goes on and on. While a daughter language evolve from a parent, it accumulates words from other languages it is in contact with at that period, but its basic grammar remains the same as its parent. Linguist use these grammatical constructions as one of their main tools to trace related languages. There are occasions when one language dominates another, progressively invades it and replaces it altogether. This process is called *Elite Dominance*.

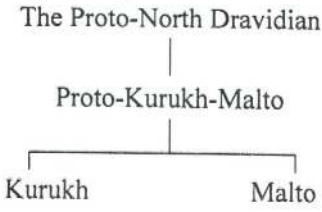
The Tamil language too had an early parent language which we call *Proto-South Dravidian*; an earlier grand parent language called *Proto-Dravidian* and an earliest great-grand parent language hypothetically called *Proto-Elamo Dravidian*. There would be still earlier languages, but we do not know anything about them at present. The evidence to show that the Tamil has had a great-grand parent language which gave rise to the Elamite and Sumerian languages in the Middle East and to the Proto-Dravidian in South Asia comes mainly from the genetic evidence corroborated by the linguistic studies.

The Proto-Elamo Dravidian languages were spoken in Iran, the Fertile Crescent (which included South-East Iraq) and South Asia about 9,000 years ago. Elamite, and possibly Sumerian, two of the oldest languages spoken by man that we are aware of were off-springs of this family (Sathasivam A. 1965; McAlpin, D. 1981). Sumerian is perhaps the first language in the world to develop a script. The Sumerian clay tablets with their cuneiform writings have been dated by radiometric methods to 4500-2300 BCE (years Before Common Era). About 3,500 years ago the Sumerian language became extinct.

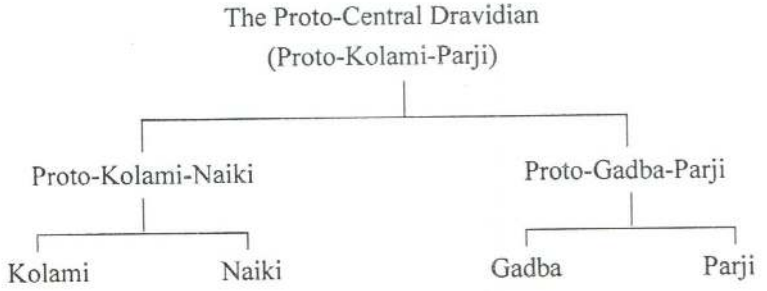
Linguists have worked out the steps through which the modern Dravidian languages have evolved from the Proto-Dravidian. In Baluchistan, the Proto-Dravidian gave rise to a Proto-Brahui, which gave rise to the Brahui language. Tribes speaking the Brahui language are found in Iran, Afghanistan and Pakistan. In India the off-springs of the Proto-Dravidian are divisible into the Proto-North Dravidian, the Proto-Central Dravidian, the Proto-South Central Dravidian and the Proto-South Dravidian.



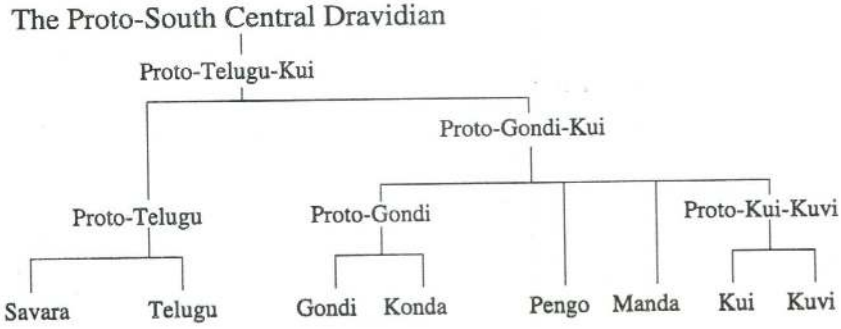
The **Proto-North Dravidian** gave rise to a Proto-Kurukh-Malto, from which evolved the Kurukh and Malto Languages. During the early 21st century Kurukh was spoken by about 1.75 million people predominantly among the Oraon tribes of Chota Nagpur and in parts of West Bengal and Bangladesh. In many parts of North India it has been displaced by the Hindi language. Kurukh and Malto share several articulated sound types with Brahui, suggesting a common undivided stage in the ancient past.



The **Proto-Central Dravidian** is also called Proto-Kolami-Parji. From this evolved two language groups Proto-Kolami-Naiki and Proto-Gadba-Parji. The former gave rise to Kolami and Naiki, while from the latter evolved Gadba and Parji languages of Central India.



The **Proto - South Central Dravidian** is named Proto-Telugu-Kui. This gave rise to two language groups: Proto-Telugu and Proto-Gondi-Kui. The Proto-Gondi-Kui branched out into Proto-Gondi, Proto-Kui-Kuvi groups and the languages Pengo and Manda. The Proto-Gondi gave rise to Gondi and Konda languages, while the Proto-Kui-Kuvi gave rise to Kui and Kuvi. The Proto-Telugu gave rise to the Savara and Telugu languages.

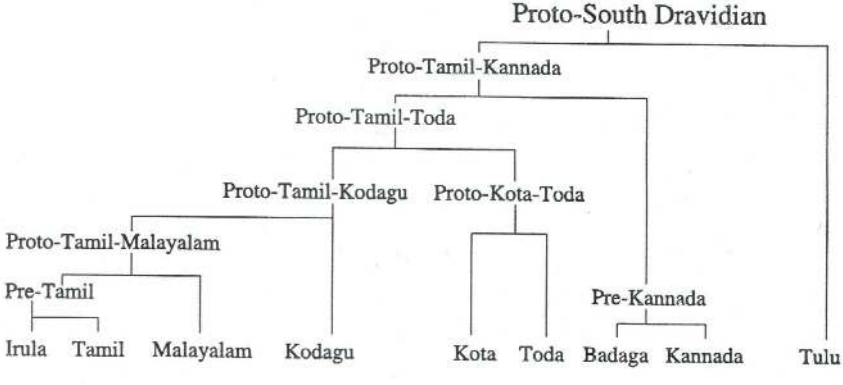


The **Proto-South Dravidian** branch of the Proto-Dravidian language gave rise to the Proto-Tamil-Kannada group and the Tulu language. The Proto-Tamil-Kannada gave rise to the Proto-Tamil-Toda and the Pre-Kannada languages. The Proto-Tamil-Toda gave rise to Kota and Toda, while the Pre-Kannada gave rise to the Badaga and Kannada languages.

The Proto-Tamil-Kodagu divided into the Proto-Tamil-Malayalam and the Kodagu language. The Proto-Tamil- Malayalam gave rise to a Pre-Tamil and Malayalam languages. The Irula language branched off from Tamil.

Linguistic reconstruction suggests that Proto-Dravidian was spoken in Northern India in the Indus and Gangetic regions during the fifth millennium BCE associated with the Neolithic culture. The reconstructed vocabulary for Proto-Dravidian includes terms dealing with agriculture, herding and hunting.

During the third millennium BCE the Proto-Dravidian was spoken in the region around the lower Godavari basin in peninsular India. The evidence relating to terms dealing with agriculture, irrigation and herding, economic activity, social strata and religion including words relating to deities indicate that the speakers of Proto-Dravidian were responsible for the culture associated with the Neolithic Complexes of South India.



Stages in the Evolution of Tamil from the Proto-South Dravidian

Stages in the Evolution of Tamil from the Proto-South Dravidian

The reconstructed terminology for Proto-South Dravidian includes urban settlements and buildings such as palaces, prisons, administrative and military structures, social strata and castes and an advanced metal-based technology. The linguistic evidence suggests that Proto-South Dravidian was spoken around the middle of the second millennium BCE, but the earliest archaeologically attested community that displays the above traits is the Tamil polity of the Sangam Period, which dates to the latter-half of the first millennium BCE.

The earliest form of Tamil speech (Proto-Tamil-Malayalam) probably existed by about 1500 BCE (Kamil Zvelebil 2003: 698). The exact period when the name 'Tamil' came to be applied to the language is unclear, but Tamil language written with Brahmi characters in potsherds has been dated to around 400 BCE (Rajan, K. 2009: 57-84; Mahadevan, I. 1995: 173-188).

In Amaravati, in present day Andhra Pradesh there is an inscription referring to a *Dhamila-vaniya* (Tamil trader) datable to the third century BCE. Another inscription from Nagarjunakonda also in Andhra Pradesh seems to refer to Damila. A third inscription from Kanheri refers to a *Dhamila-gharini* (Tamil householder) (Indrapala, K. 2006: 156). The *Hathigumpha Inscription* of the Kalinga ruler Kharavela datable to 150 BCE refers to a *Tramira samghata* or Confederacy of Tamil rulers that had been in existence for 113 years before the date of the inscription (Epigraphia Indica XX, 5: 79-86). Thus, by about 300 BCE, if not before, not only the ethnic identity of the Tamils as a distinct group, but there existed an organised government to rule themselves in South India is clearly established.

During the Sangam Period the country ruled by the Cheras was a Tamil speaking land. It was the Chera country which produced the great Tamil epic *Silappadikaram*. Malayalam is believed to have separated from Tamil by the fifth century. The present day Malayalam language has more than 50% of Sanskrit words in its vocabulary.

Genetic Variation among the Dravidians

The Dravidians unlike some other linguistic groups do not belong to a single racial or genetic group. They originated from different genetic backgrounds but came to live together and to speak a particular family of languages in South Asia forming a specific linguistic group. Y-Chromosome DNA analysis of the South Indian Dravidian peoples including the Tamils conducted by several investigators over the past ten years under the Genographic Project, as well as other programmes have shown that there are about 7 major and several minor genetic groups showing specific genetic markers among the present-day Dravidian speakers. The most important ones are:

1. Haplogroup C (Australoid group – Y Chromosome DNA marker *M130*): These genetic markers are present among 5% of the males in South India. The earliest carriers of this genetic marker arrived in South India 60,000 – 55,000 years ago.

2. Haplogroup F: (Middle Eastern group – Y Chromosome DNA marker *M89*): This genetic marker evolved from the first group of *Homo sapiens* to come out of Africa. It is present at a frequency of 18.1% among the Tribal groups and 9.6% among the castes; its overall frequency being 12.5%.

3. Haplogroup H (South Asian group – DNA marker *M69*): This group was part of the major inland settlement of India. Geneticists believe that this group originated somewhere along the migration route of peoples carrying the *M20* chromosome marker arriving in South India 35,000 years ago. Today this group is present among 25-40% of the Dravidian speaking people.

4. Haplogroup L (Indo-Mesopotamian group – DNA marker *M20*): This genetic marker and its sub-groups are present among 20-30% of the males in South India and their skull type is described as *Dolicocephalic* or *Mediterranean Dravidic* type. This genetic group first arrived in South India 30,000 years ago from the Middle East through a northern route.

5. Haplogroup R2 (Central Asian group – DNA marker *M124*): This group arrived from South Central Asia 25,000 years ago. This is present among 5% of South Indians.

6. Haplogroup J2 (Mesopotamian group – DNA marker *M172*): This group arrived in India from Elam in Mesopotamia 10,000 – 9,000 years ago. This is present in 15% of the males in Tamil Nadu. They too represent the Mediterranean skeletal type.

7. Haplogroup R1A1 (Indo-Aryan group – DNA marker *M17*): This group Arrived in North India from Central Asia via Iran 3,800 years ago and have spread far and wide all over South Asia. Major sections of them arrived in South India around 500 BCE, about 1,000 years after the onset of the South Indian Megalithic Culture. They now represent about 11% of the Tamil Nadu, 18% of the Kerala and 13% of the Sri Lanka male populations.

Among the minor groups: Haplogroup O3 (Chinese group – DNA marker *M122*) arrived in India from China 6,000 – 4,000 years ago. This group is believed to be responsible for the introduction of rice into India. They mainly occupy the sub-Himalayan regions, and a high percentage is found in Bengal. Overall their presence is about 1% in South India.

Equating with the earlier anthropological classification Haplogroup C, and possibly F and H represent the Australoid or Austro-Asiatic group; Haplogroups L, R2, J2 and R1a1 represent the Caucasoids and O3 represents the Mongoloids.

Thus, the Caucasoid, Australoid and Mongoloid groups and various types of combinations of these groups are found among the Dravidian speaking peoples of South India. There is no single Dravidian race or a specific anthropological group which could be named Dravidian.

It is also relevant to note that more than 70% of the present-day speakers of the Dravidian languages are the direct descendants of the ancestral populations that have occupied South Asia more than 30,000 years ago.

Bibliography

BURROW, THOMAS (1960a): Sanskrit and the Pre-Aryan Tribes and Languages, *Indo- Asian Culture*, April 1960.

BURROW, THOMAS (1960b): The Sanskrit Language, Faber and Faber, London.

CAMPBELL, A.D. (1849): A grammar of the Teloogoo language, peculiar to the Hindoos occupying the northeastern provinces of the Indian peninsula, 3rd edition, Hindu Press, Madras.

CALDWELL, ROBERT (1856); A Comparative Grammar of the Dravidian or South- ndian Family of Languages, Harrison, London. (repr.1956):University of Madras.

CHATTERJI, S.K. (1965): Dravidian, Annamalai University, Madras.

INDRAPALA, K. (2005): The Evolution of an Ethnic Identity, Sydney; Reprint (2006), Madras.

KRISHNAMURTHI, BADRIRAJU (2003): Dravidian Languages, Cambridge University Press.

MAHADEVAN, IRAVATHAM (2003): Early Tamil Epigraphy: From the Earliest Times to the Sixth Century A.D., *Harvard Oriental Series* 62, Chennai.

Mc ALPIN DAVID (1981): Proto-Elamo-Dravidian: The Evidence and its Implications, American Philosophical Society, Philadelphia.

RAJAN, K. (2009): *Damili Graffiti and Cave Records* in INDRAPLA, K. (ed.) Early Historic Tamil Nadu, Kumaran Book House, Chennai.

SATHASIVAM, A. (1965): Sumerian – A Dravidian Language, Berkley, USA.

ZVELEBIL, KAMIL (2003): Dravidian Languages, New Encyclopaedia Britannica Britannica, 22.

ZVELEBIL, KAMIL (1990): Dravidian Linguistics: An Introduction, Pondicherry Institute of Linguistics and Culture, Pondicherry.

Cover Illustration by
S.Gauthaman

This monograph about Tamils, their genetic origins and how their language evolved from an early proto-Dravidian which have a relationship with the proto-Elamite and proto-Sumerian languages of Mesopotamia are presented here in three academic papers.

Dr Siva Thiagarajah is a medical doctor with a special interest in Human Genetics, Archaeology and History; a well known writer and an academic researcher. He is a Biological Science graduate from the University of Ceylon; a Medical graduate from the Faculty of Medicine, University of Ceylon and a Doctor of Philosophy from the University of London for his work in Genetics.

In 1980 he was awarded the prestigious Doctor Green Memorial Prize by the Jaffna Medical Association for his Four-year Study of Ectopic Pregnancy at the Jaffna General Hospital 1977-1980 (now, the University teaching hospital).

To-date he has published thirteen books including a popular book on Medicine in Tamil; a Tamil translation of The Epic of Gilgamesh, the world's first book ever written in clay tablets 5,000 years ago in the lost Sumerian language and brought to light through twentieth century archaeology; a history book about the religious beliefs among the Tamils over the past 5,000 years named Tamils and their Religions; and a Study based on Population Genetics and Archaeology, Peoples and Cultures of Early Sri Lanka, narrating the story of the human migrations in South Asia over the past 55,000 years, acclaimed as a scientific tour-de-force.

Kumaran Book House

Subject: History

ISBN 978-955-659-473-7

9 789556 594737

Price Rs. 450.00

660