Chapter 3:

Materials and Methods

3. Materials and Methods

The details about the techniques, analytical parameters and the statistical tools that were used for collection, analyses and interpretation of the data have been discussed in this chapter. It has been subdivided into three sections. In the first section, the criteria and the procedures that were followed for the collection of the samples have been discussed. In the second section, the experimental procedures that were carried out in the laboratory have been elaborated. Finally, in the third section, the applications of the statistical tools for our data analyses have been discussed. In order to facilitate better understanding, each of these three sections has been subdivided into two subcategories: Population based study and Rheumatoid Arthritis (RA) based study.

3.1. Sample Collection

This section has been categorized into two parts. In the first part, sample collecting procedure from different populations, have been discussed while the second part contains the description of sample collection in RA patients and control groups.

3.1.1. Population Based Study

3.1.1.1. Study Populations

In this part of the study, five populations were selected from Northern part of West Bengal with the aim of assessing and quantifying the KIR genetic variation existing within and between the populations. Populations that were included in the study were not only selected from scheduled caste and scheduled tribe ethnicity but also included individuals from community based sampling. Randomly selected normal healthy individuals from each population have been included in the present study. Three major selection criteria were considered for sample collection, which includes:

[1] Only individuals unrelated from each other for the last three generations were selected for the study.

[2] Geographical strictness was maintained while collecting samples from a population. It was made sure to collect all the samples of a particular population from the geographical locations as mentioned in our study.

[3] Individuals having inter-caste or inter-religion marriages, in any of the last three generations, were excluded from the analyses.

The five populations that were selected for the study include:

1. Rajbanshi: an ethnic caste population from the Terai and Dooars region of West Bengal,
2. Bengali: a heterogenous population group from Siliguri and adjoining regions of West Bengal,
3. Rabha, a primitive tribal population group from the Northern part of West Bengal,
4. Gurkha, a Nepali speaking ethnic population from Siliguri and adjoining areas of West Bengal, and
5. Muslim, from Northern part of the state

Detailed description about the possible origin, socio-cultural practices and the position of these population groups in the social hierarchy network of Indian society have been discussed in detail below.

3.1.1.1.1. Rajbanshi

The Rajbanshis represent a highly diversified ethnic community of India. They have rich cultural, linguistic and social background. They account for 18.4% of the total Scheduled Caste population of West Bengal (Census of India, 2001). Although distributed dispersedly throughout the state, the Rajbanshis mainly inhabit the Terai and Dooars of the Northern region of West Bengal, especially the districts of Jalpaiguri and Coochbehar. They can also be found adequately in the bordering state of Assam. They have an Indo-European linguistic background. Besides their own dialect, they also speak

Bengali, Assamese and some other minor languages. They practice Hinduism and have their own rich cultural identity and heritage. Earlier reports suggested that the Koch Rajbanshis are represented by many sub-castes in North Bengal, India (Risley, 1891). Thus, it can be said that the ethnicity and genetic richness of the Rajbanshis may enhance their importance in genetic diversity studies and interpretation of the migratory routes in this part of the country.

3.1.1.1.2. Bengali

Many communities in India have emerged because of the blending of different ethnic populations. Such a community is the Bengali community. They are the major inhabitants of the state of West Bengal and are the third largest ethnic group in the world. In West Bengal, the Bengalis mainly follow Hinduism (Census of India, 2001). Many of them also follow Islam, with considerable traces of Christianity and Buddhism (Census of India, 2001). They speak the Bengali language or the 'Bangla' dialect, which is thought to belong to Indo-European group. They are culturally very rich and this made them as one of the unique population groups of India. Since the Bengali community is considered to have a mixed descendant, there is an urgent need to unravel the genetic profile of this extremely heterogenous population, beside exploring the genetic diversity of indigenous tribes and castes.

3.1.1.1.3. Rabha

Rabha is a very little known small endogamous Scheduled Tribe community of India (Census, 2001;Chakraborty, 2013;Sarkar, 2011) with a conserved gene pool of their own. In West Bengal, they mainly dwell in the forest villages of the Dooars region of Coochbehar and Jalpaiguri districts. Historically, they are considered as the primitive inhabitants of the region who remained isolated from other neighbouring populations due to strict endogamy. According to H.H. Risley, the Rabhas belong to the Indo-mongoloid stock (Risley, 1891), having a unique genetic constitution. Thus, there is an urgency to undertake genetic diversity studies in this group.

3.1.1.1.4. Gurkha

The Indian Gurkhas constitute a community of Nepali speaking people, populating the Eastern and North-Eastern states of the country, mainly in West Bengal and Sikkim (Chatterjee, 1974) with sizeable populations in the states of Meghalaya, Nagaland, Manipur, Tripura, Mizoram, Arunachal Pradesh as well as in Assam. In West Bengal, they are distributed in the Terai and Dooars, as well as in the hilly terrains of the Northern part of the state. Apart from India, they are found in Nepal and to some extent in Bhutan. They are a mixture of various clans and ethnic groups which include Bahun (Brahmins), Chhetri, Thakuri, Gurung, Magar, Newar, Tamang, Rai (Khambu), Limbu (Subba), Sherpa, Yolmo and many more (Roy, 2012). Each of them has their own languages, but Nepali language has become the common binding thread for all the Gurkha castes and clans.

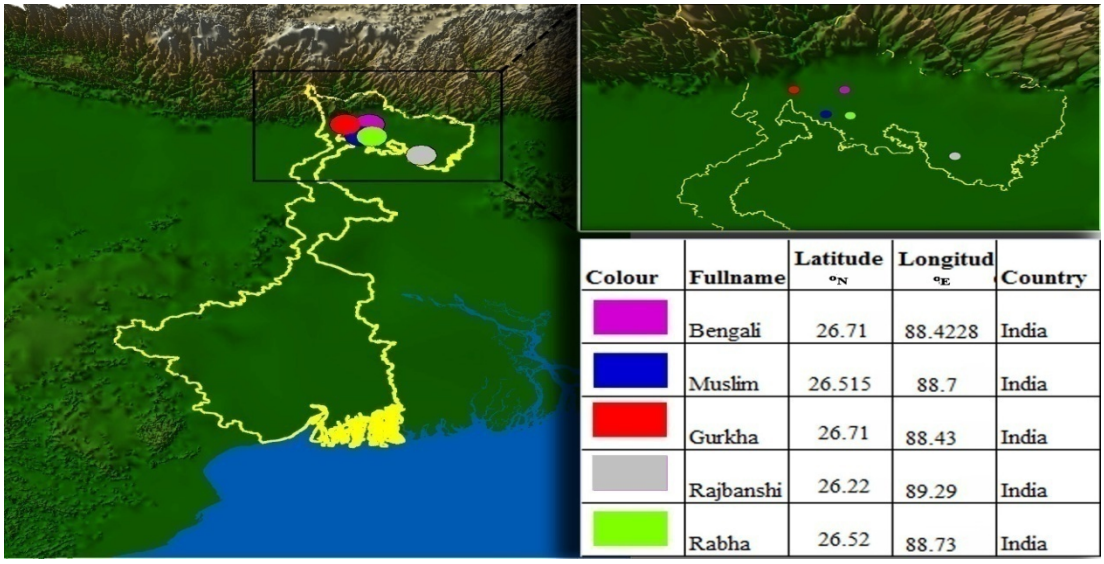
3.1.1.1.5. Muslim

The Bengali Muslims inhabiting West Bengal represent the second-largest ethnic Muslim community in the world, after the Arabs (Eaton, 1993), native to modern-day Bangladesh and the Eastern states of India, including West Bengal and Assam. They speak Bengali dialects and have strong cultural similarities with the Bengali Hindus, thereby increasing the cultural richness of West Bengal. Being the second-largest community and also the largest minority group of the state, the Bengali Muslims comprise 27.01% of the total population of West Bengal (Census of India, 2011).

3.1.1.2. Sample collection strategy

Prior to sample collection, geographical locations/sites for collecting the samples were determined based on the demographic history of the place. On the day of sample collection, volunteers were questioned about their ethnicity, caste and tribal affiliations and surnames and the birthplaces of their parents. The demographic profile and other ethnical and familial information were filled in a detailed questionnaire **(Appendix 1)**. Based on the questionnaire, only unrelated subjects were considered eligible to participate in the study. Three- generation pedigree charts were prepared to ensure un-

relatedness in all the samples. Two (2) ml of whole blood was obtained from the volunteers with their informed consent by venipuncture method and was collected in EDTA vacutainer tube. The blood samples were stored at -20°C until use. The investigation was approved by the Human Ethics Committee of the Department of Zoology, University of North Bengal, India. The different sample collecting sites in case of the population-based study are shown in Figure 20.



Colour	Fullname	Latitude °N	Longitud °E	Country
	Bengali	26.71	88.4228	India
	Muslim	26.515	88.7	India
	Gurkha	26.71	88.43	India
	Rajbanshi	26.22	89.29	India
	Rabha	26.52	88.73	India

Figure 20: Geographical Location of the different population group in the map of west Bengal.

3.1.2. RA Based Study

3.1.2.1. Study population

A total number of 110 Rheumatoid Arthritis patients were included in this study. The samples were collected from an authorized diagnostic laboratory of Siliguri and from North Bengal Medical College and Hospital (NBMCH, Sushrutnagar, Siliguri, West Bengal, India) under the guidance of medical practitioner. The diagnosis of RA was made by the physician based on the medical and clinical history, physical examinations and symptoms of the disease and most importantly, their fulfilment of the American College

of Rheumatology criteria 2010 (Aletaha, *et al.*, 2010). The inclusion criteria for the subjects in the patient group for the study included the following:

1. Resident of Siliguri and adjoining areas of the sub-Himalayan region of West Bengal, India.
2. Subjects fulfilling the ACR criteria 2010.
3. Subjects having the symptoms for ≥1 year duration.
4. Patient should be above 18 years of age.
5. Patients capable of giving informed consent for becoming a part of the study.

According to ACR criteria 2010 (Aletaha, *et al.*, 2010), patient is considered to have definite RA, if he/she has 1 or more swollen joint(s) and having ≥6 score on the classification criteria which include joint distribution, serology reports, disease duration and measures of acute phase reactants. Furthermore, evaluation of the disease was also done based on DAS28 score system for confirmation. Most of the patients belonged to stage II and III as was also evident from the clinical evaluations and other relevant data.

However, there are certain exclusion criteria that were followed in the study. These are listed below:

1. Patients having a history of or currently suffering from primary inflammatory joint diseases or primary autoimmune diseases other than RA.
2. Patient suffering from known HIV or hepatitis B/C infection.
3. Individuals having latent TB infection unless they have completed adequate antibiotic prophylaxis.
4. Reported cases of malignancy (other than basal cell carcinoma) within the last 10 years
5. Patients with demyelinating disease.

6. Known history of recent drug or alcohol abuse

7. Patients with poor tolerability of venipuncture or patients with lack of adequate venous access for required blood sampling during the study period.

8. Patients currently involved in other clinical trial(s) involving an investigational medicinal product.

9. Other severe acute or chronic medical or psychiatric condition, or laboratory abnormality, which according to the investigator, may impart excess risk associated with study participation, or which would render the patient inappropriate for becoming a part of the study.

3.1.2.2. Sample collection strategy

The demographic profile and other ethnical and familial information of individual volunteer were filled in a detailed questionnaire **(Appendix-1)**. Based on the questionnaire, only unrelated subjects were considered as eligible to participate in the study. Three-generation pedigree charts were prepared to ensure the un-relatedness in all the samples. The patients were also asked to complete the Health Assessment Questinnaire (HAQ) (**Appendix 2)** for their inclusion in the study. Five (5) ml of whole blood was then obtained from the volunteers with their informed consent by venipuncture method. Three (3) ml of the sample was allowed to stand for serum preparation while the remaining 2 ml was stored in EDTA vacutainer tubes for genotyping study, at -20°C, until use. Two (2) ml blood was separately collected for ESR test.

Both patients and control subjects have provided their written consent after knowing the purpose of the study. The non-RA or the control group included patients with spondyloarthropathy (n=12), osteoarthritis (n=2), gout (n=7), undifferentiated arthritis (n=13), mechanical pain (n=22) and healthy subjects (n=46). The investigation was approved by the Human Ethics Committee of the Department of Zoology, University of North Bengal, India.

3.2. Experimental Procedures:

This section is further divided into two subcategories, population-based study and RA based study, based on the methodologies that were followed. Detailed description of all the chemicals, reagents and kits (**marked in bold**) used for the study are given in **Appendix 3.**

3.2.1. Population based study

Below are the protocols that were employed for KIR genotyping of the populations that have been selected for our study.

3.2.1.1. Extraction of genomic DNA

Genomic DNA was extracted from frozen peripheral blood samples that was collected from the individual volunteers, using phenol-chloroform extraction method (Sambrook, *et al.,* 1989) with certain modifications.

- Blood samples initially kept at -20 °C, were thawed first for about 20 minutes prior to use.
- 500 µl of the thawed blood sample was taken into a 2 ml microfuge tube and washed with 1 ml of 1X SSC, mixed gently and centrifuges at 5000 rpm for 5 minutes.
- The supernatant was discarded and 1.2 ml of 1X SSC was added. The content was mixed gently and centrifuged at 5000 rpm for 3 minutes.
- The supernatant was discarded again. The pellet was resuspended with 1.2 ml of 50 mM KCl and centrifuged as above.
- After centrifugation, the supernatant was discarded again. 375 µl of High Salt Buffer was added to the pellet, followed by the addition of 25 µl of 10% SDS and 12.5 µl of 8 mg/ml Proteinase K. The content was mixed gently and incubated at 56 °C for 1 hour.
- To the Proteinase K digested suspension, equal volume of Phenol-Chloroform was added and centrifuged at 12000 rpm for 20 minutes at 4°C .
- The DNA was precipitated in chilled absolute alcohol.

- ❖ Then, the DNA was rinsed twice with 500 µl of 70% ethanol.
- ❖ The DNA was then, dried and dissolved in 50 µl **TE buffer**.

3.2.1.2. Characterization of DNA

3.2.1.2.1. The integrity:

The integrity of the extracted DNA is an important factor to consider during extraction steps. Integrity was checked by electrophoresis on 1% agarose gel prepared in 1X TBE buffer (**Appendix 4**), containing Ethidium bromide (3l of 10mg/ml stock for every 50 ml of 1% agarose). The high molecular weight genomic DNA appeared as a single band near the well (**Figure 21**).

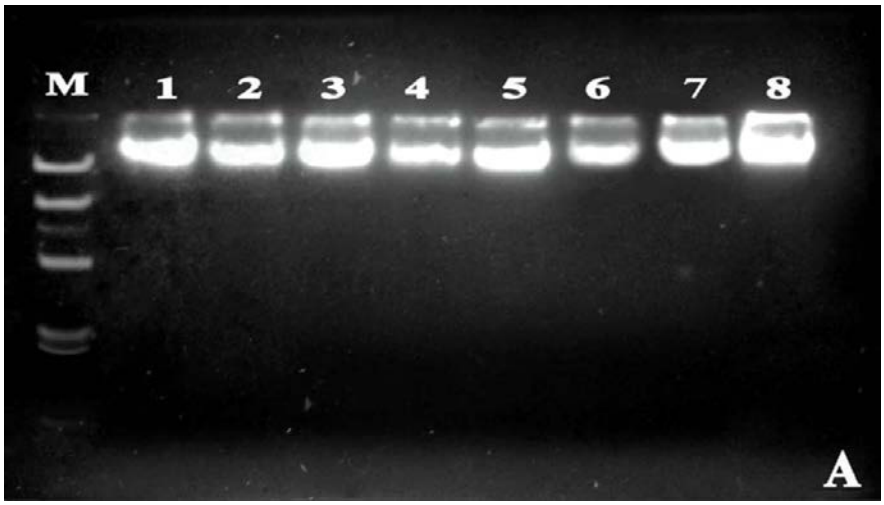


Figure 21:1% agarose gel electrophoresis to demonstrate the extracted DNA in fresh bloods. Lane M: M stands for Marker (λ-DNA/Hind III Marker).Lane 1-8: DNA from Rabha population (sample 1-8).

3.2.1.2.2. The Purity and concentration:

Five (5) µl of stock genomic DNA was diluted to 1.0 ml using deionized water (dilution factor = 200), mixed well and optical density was measured at 260 nm and 280 nm in a UV spectrophotometer (Rayleigh UV- 2100, China). DNA samples with 260/280 ratio greater than 1.7 but less than 2.0 were considered to be free from protein contaminations.

In the cases, where the ratio was less than 1.2, DNA was extracted again. Each DNA sample was checked on ethidium bromide prestained agarose gel by comparing with known quantity of phage-Lambda DNA as estimation by absorbance may be inaccurate due to RNA contamination. Samples, with large amount of RNA contamination were treated with RNAase A (Sigma Aldrich, final concentration of 100g /ml) at 37 °C for 2 hrs, followed by re-extraction with phenol-chloroform iso-amyl alcohol mixture.

DNA concentration of the sample was calculated using the formula:

OD at 260 nm X DF X 50 (1 OD = 50μg / ml of ds DNA)

The original DNA solution was diluted by a factor, DF = 200

E.g. if OD of the diluted DNA sample = 0.016

DNA concentration = 0.016 X 200 X 50 μg / ml = 160 μg / ml

3.2.1.3. Storage

Samples which were in regular use were kept at 4 °C. Rest of the samples were stored at -80°C.

3.2.1.4. PCR-SSP Typing

Molecular typing of the genomic DNA was performed for detecting the presence of the following KIR genes : *2DL1, 2DL2, 2DL3, 2DL4, 2DL5, 2DS1, 2DS2, 2DS3, 2DS4, 2DS5, 3DL1, 2DL2, 3DL3, 3DS1, 2DP1 and 3DP1* using 16 PCR–sequence-specific priming (PCR-SSP) reactions.

3.2.1.4.1. PCR Primers

The primer sequences were generously provided by R. Rajalingam (Immunogenetics and Transplantation Laboratory, University of California) on query (Rajalingam, 2011;Rajalingam, *et al.*, 2008). Each primer carries a 3′ residue matching a unique position conserved on all known sequences of a given KIR gene (Garcia, *et al.*, 2003).

Primer lengths were adjusted in such a way so that the annealing temperatures are maintained between 59°C and 67°C, which enables amplification of all KIR genes under the same PCR condition. In addition to gene-specific primers, internal positive control primers specific to the conserved adenomatous polyposis gene were included to confirm each PCR reaction. The oligonucleotide primers used in the KIR genotyping system are provided in Table 5.

Table 5: Primer sequences and the amplicon sizes of the 16 KIR genes.

Target	Forward primer		Reverse primer		Genomic amplicon (bp)
	Name	Sequence (5'-3')	Name	Sequence (5'-3')	
KIR genes					
2DL1	2DL1F	CCATCAGTCGCATGACG	2DL1R1 2DL1R2	CCACTCGTATGGAGAGTCAT AATGTTCCGTTGACCTTGGT	1903 & 1818
2DL2	2DL2F2	ACTTCCTTCTGCACA(C/G)AGAA	2DL2R1	CCCTGCAGAGAACCTACA	1877
2DL3	2DL3F3	CTTCATCGCTGGTGCTG	2DL3R1	CAGGAGACAACTTTGGATCA	816
2DL4	2DL4F1	CTGCATGCTGTGATTAGGTA	2DL4R1	CTGTTGAGGGTCTCTTGCT	695
2DL5	2DL5F	TGCCTCGAGGAGGACAT	2DL5R1	TCATAGGGTGAGTCATGGAG	1151
3DL1	3DL1F1	AT(C/T)GGTCCCATGATGCT	3DL1R1	CTGAGAGAGAAGGTTTCTCATATG	1661
3DL2	3DL2F1	TGCAGGAACCTACAGATGTTAT	3DL2R1	CTTGAGTTTGACCACACGC	1882
3DL3	3DL3F1	CACTGTGGTGTCTGAAGGAC	3DL3R1	TCTCTGTGCAGAAGGAAGC	1905
3DS1	3DS1F	GGCAGAATATTCCAGGAGG	3DS1R1	GGCACGCATCATGGA	1847
2DS1	2DS1F1 2DS1F2	CTCCATCAGTCGCATGAG CTCCATCAGTCGCATGAA	2DS1R	AGGGCCCAGAGGAAAGTT	1922 & 1897
2DS2	2DS2F	TGCACAGAGAGGGGAAGTA	2DS2R1	CGCTCTCTCCTGCCAA	1781
2DS3	2DS3F	TCACTCCCCCTATCAGTTT	2DS3R	GCATCTGTAGGTTCCTCCT	1812
2DS4	2DS4F1	TCCTGCAATGTTGGTCG	2DS4R1	ACGGAAACAAGCAGTGGA	2050
2DS4-full	2DS4F1	TCCTGCAATGTTGGTCG	2DS4fullR1	CCCTCCCTGGATAGATGGTAC	1956
2DS4-del	2DS4F1	TCCTGCAATGTTGGTCG	2DS4delR	TTCCCTGGATAGATGGAGCTG	1933
2DS5	2DS5F	AGAGAGGGGACGTTTAACC	2DS5R	GGAAAGAGCCGAAGCATC	1952
	2DS5F	AGAGAGGGGACGTTTAACC	2DS5RD	CAGAGGGTCACTGGGC	180
2DP1	2DP1F	TCTGTTACTCACTCCCCCA	2DP1R	GGAAAGAGCCGAAGCATC	1825
3DP1	3DP1F1	AGAGTATTCCGAAACACCG	3DP1R1	CTGACAACTGATAGGGGGAA	1900
Internal positive controls					
IC256	PIC-F	ATGATGTTGACCTTTCCAGGG	PIC-R	ATTGTGTAACTTTTCATCAGTTGC	256
GH1	GH1F	CTTCCCAACCATTCCCTTA	GH1R	CGGATTTCTGTTGTGTTTC	450

If the primer is marked as xx.xx nmol, then the amount of MilliQ (MQ) water that is to be added to the primer solution to make its concentration 100 µM is 10 X (xx.xx) µl. For example, if the primer is marked in the tube or product sheet as 59.68 nmol, one needs to add 596.8 µl MQ water to make this primer as 100 µM. This becomes the stock solution which can then be stored at -20 °C until use. In case of an ideal PCR, the working primer

concentration is 10 µM. In order to prepare this, 50 µl of 100 µM primer stock was added to 450 µl water. This can be stored at 4 °C for 3-4 months.

3.2.1.4.2. Preparation of reaction mixture:

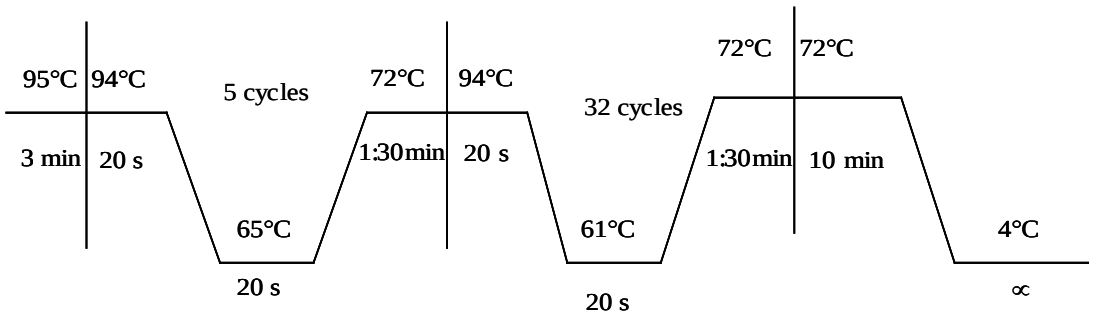
Reaction mixtures of 25 µl volume were used to carry out the PCRs using a thermal cycler (MiniTM Gradient Thermal Cycler, PTC-1148, Bio-Rad, Singapore). The reaction mix of 25 µl was prepared in the following proportions:

PCR mixture	initial concentration	vol(µl)/25ul	final concentration.
MilliQ H2O	1X	6.38	1X
10x PCR buffer	100 mM	2.5 µl	10 mM Tris, 50 mM KCl
dNTP mix	2.5 mM	2 µl	200 µM each
KIRF (KIR-specific)	10 µM	1.5 µl	0.6 µM
KIRR (KIR-specific)	10 µM	1.5 µl	0.6 µM
Internal Control (Forward)	10 µM	1 µl	0.4 µM
Internal Control (Reverse)	10 µM	1 µl	0.4 µM
AmpliTaq polymerase	5 U/ul	0.2 µl	1U/25 µl
DNA	100 ng/µl	2.0 µl	8 ng/µl
Total volume		25 µl	

3.2.1.4.3. PCR typing protocol

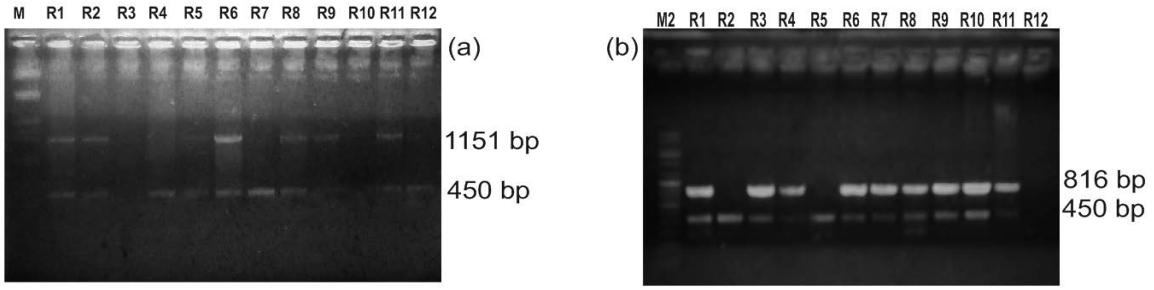
PCR protocol included an initial denaturation step at 94°C for 3 min, then five cycles of 94°C for 30s, 62°C for 50s, and 72°C for 1 minute. This was followed by 30 cycles of 95°C for 30 s, 60°C for 50s, and 72°C for 1 minute with small modifications of the annealing temperature for different primer sets (shown below). The numbers of amplification cycles were modified to 35 cycles for KIR2DS1 and 2DS4 and 32 cycles

for KIR2DS2 and 2DS5. The number of cycles for the second segment were adjusted between 30-35cycles for activating KIR genes, if needed.



3.2.1.4.4. Gel checking

Appropriate aliquotes of PCR products were analyzed on ethidium bromide (0.5 µg/ml) prestained 2% agarose gels. TBE buffer was used as tank buffer, Bromophenol blue as the tracking dye and 1Kb plus DNA ladder (0.1- 12 Kb) (Catalog number: 10787018: Thermo Fisher Scientific, Massachusetts, U.S.A) was used as DNA size marker. After electrophoresis, a photograph of the agarose gel was taken over a UV-transilluminator. This was followed by careful examination of the gel pictures to determine positive amplifications against appropriate molecular weight markers (Figure 22). Each lane of the gel except for a negative control lane, are loaded with PCR product and showed a control band. Sixteen KIR gene were typed on a selected DNA sample. In case of false reactions where no control bands are found, the reactions were repeated.



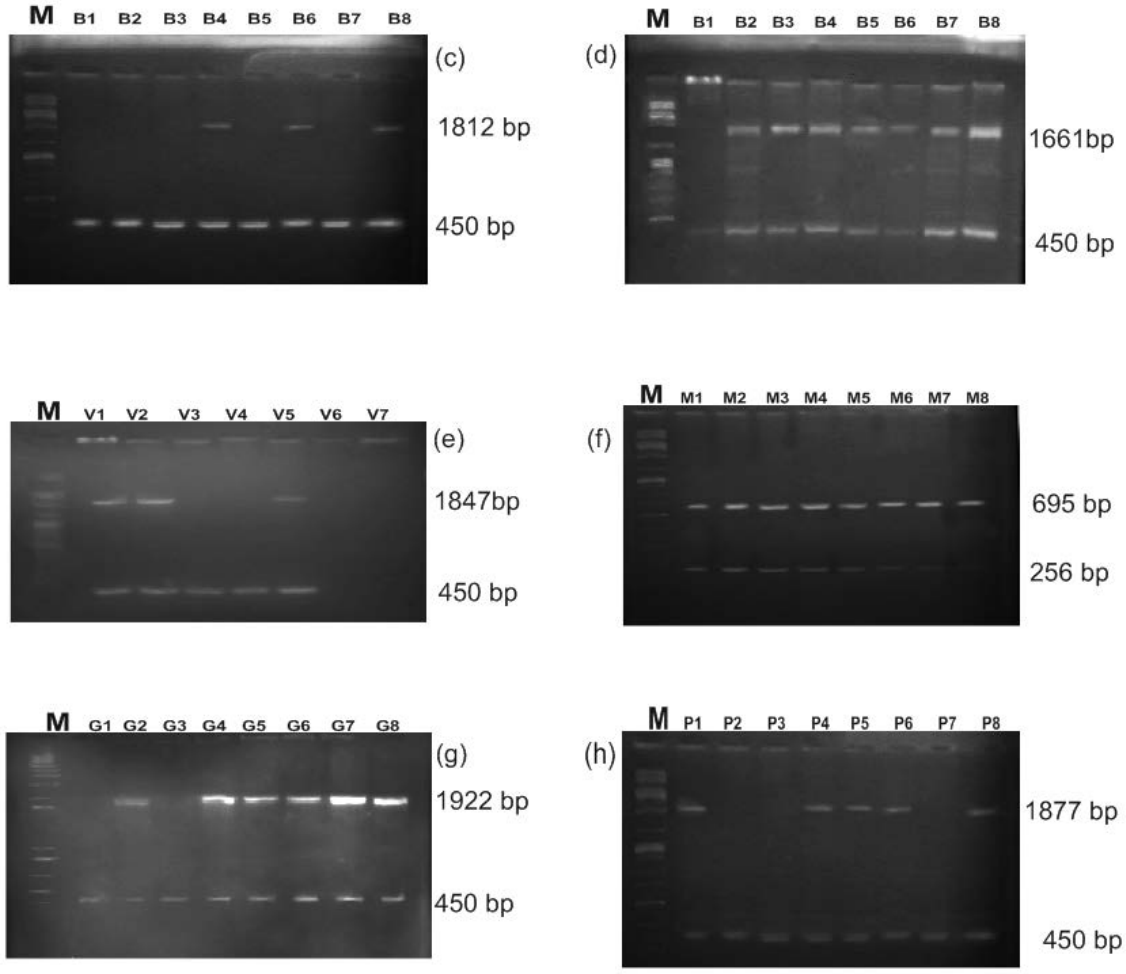


Figure 22: 2% agarose gel electrophoresis to show the presence of different KIR genes. [a] KIR2DL5 (1151 bp) in the Rajbanshi population [RAJ10-21 in lane R1-12]. [b] KIR2DL3 (816 bp) in the Rajbanshi population [RAJ10-21 in lane R1-12]. [c] & [d] KIR2DS3 (1812 bp) and KIR3DL1 (1661 bp) respectively in the Bengali population [BEN25-32 in lane B1-B8]. [e] KIR3DS1 (1847 bp) in the Rabha population [R6-10 in lane V1-5]. [f] KIR2DL4 (695 bp) in Muslim population [M1-8 in lane M1-8]. [g] KIR2DS1 (1922 bp) in Gurkha population [G13-20 in lane G1-8]. [h] KIR2DL2 (1877 bp) in RA patients (RA25-32 in labe P1-8]. Internal control for [a,b,c,d,e,g &h] is GH1 (450 bp) and for [f] is IC256 (256bp). The marker used in Lane M is 1Kb Plus DNA Ladder.

3.2.2. RA based study

3.2.2.1. Molecular Typing of KIR genes

Molecular typing of 16 KIR genes were carried out using PCR-SSP in both the patient and the control group in order to study the association of KIR genes with Rheumatoid Arthritis, if any. We followed the same procedure as has been described under section 3.2.1.4.

3.2.2.2. RF titre assay

The assay is based on agglutination reaction with latex enhancement. In this immunoturbidimetric assay, heat inactivated IgG (antigen) with bound latex reacts with the RF antibodies in the sample and form antigen antibody complexes leading to agglutination reactions which are then, measured turbidimetrically. The Assay was performed with COBAS INTEGRA Rheumatoid factors II assay kit (Cat. No. 20764574 322) using COBAS INTEGRA 400 plus analyzer (Roche, Germany). Manufacturers' instructions were followed for quantitative estimation of RA. The lower detection limit of the test was 10.00 IU/ml. The measuring range of the assay is 10-130 IU/ml. However, samples with higher values of RA were measured using post dilution rerun. The normal reference cut-off value for RF titre in the serum was considered to be up to 20 IU/ml as per the laboratory standardizations and guidelines.

3.2.2.3. Anti-CCP estimation

The estimation of anti-CCP in the serum samples were carried out using the commercially available Elecsys anti-CCP assay kit. The assay was performed on the Cobas e 411 Analyzer (Roche Diagnostics, Mannheim, Germany). We acknowledge our association with 3gen Diagnostics Pvt. Ltd., Siliguri, India, for providing the access to the instrument and in sharing their expertise in executing the assay. The Elecsys Anti-CCP immunoassay is a two-step IgG-capture test principle immunoassay with streptavidin-coated microparticles and electrochemiluminescence detection. The procedure was

conducted according to the manufacturer's recommendations. The unit for measurement is U/mL and the results were considered positive at a cut off value of ≥17 U/ml.

3.2.2.4. C-reactive protein (CRP) estimation

Quantitative estimation of C-reactive protein (CRP) in the serum samples using particle enhanced turbidimetric assay was also performed. The test was performed with COBAS INTEGRA Cardiac C-Reactive protein (Latex) High Sensitive (CRPHS) Cobas C pack using COBAS INTEGRA 400 plus analyzer. Human CRP agglutinates with monoclonal anti-CRP antibody precoated latex particle. The precipitate is determined at 552 nm using turbidimetric analyses. The assay was performed as per the manufacturer's instruction.

3.2.2.5. Anti-streptolysin O (ASO) titre assay

In vitro quantitative immunological determinations of Anti-streptolysin in the serum samples of the patient and the control groups were done by immunoturbidimetric assay using COBAS INTEGRA Antistreptolysin O pack in COBAS INTEGRA 400 Plus analyzer. In this assay, human ASO antibodies agglutinate with latex particle coated with streptolysin O antigens. The precipitate is determined spectrophotometrically at 552 nm. The assay was performed following manufacturer's instruction. The measuring range of the assay is 20-800 IU/ml. However, higher concentrations were determined using 1:10 post dilution rerun. The normal cut off value is 200 IU/ml in adults and 150 IU/ml in children.

3.2.2.6. Erythrocyte Sedimentation Rate (ESR) Estimation

Erythrocyte Sedimentation Rate (ESR) is an indirect measure of the acute phase reaction. Furthermore, it is a simple and inexpensive measure of inflammation, estimated by Westergrens method (Westergren, 1957;Westergren, 1926) by measuring the rate at which anticoagulated erythrocytes settle down from a fixed point in an upright calibrated tube in 1 hour. Its upper limit for males and females under normal conditions are 15 mm/hr and 20 mm/hr respectively. The basic principle underlying the estimation of ESR is that the erythrocytes normally repel each other due to the net negative charges.

However, during acute phase reactions, fibrinogen and other positively charged proteins that are present in the blood increase in amount and promote rouleaux formation, which further increases the ESR.

3.2.2.6.1. Material

- Sterile syringe (2 ml)
- Sterile injection needles for drawing of blood samples.
- 3.8 % solution of sodium citrate solution as anticoagulant,
- Westergrens pipettes with a stand. Westergrens pipette is a graded glass tube divided in millimeters. It is marked from 0 (top) to 200 (bottom). The pipette should be upright with its lower end placed in the plastic container (bowl) in a vertical position and fixed with the stand using a screw.

3.2.2.6.2. Methods

- Under sterile conditions, 1 part of 3.8 % sodium citrate (anticoagulant) was mixed with 4 parts of freshly drawn venous blood in a 2 ml syringe.
- The mixture was then drawn into the Westergrens pipette upto zero (0) mark by rotating the screw.
- The tube was then set upright in a stand with a spring clip on top and rubber at bottom.
- The apparatus was left undisturbed and the reading of the lower value of the plasma column was taken at the end of each hour. This value corresponds to the ESR value of the individual.

3.2.2.7. Serum Ceruloplasmin Assay

Ceruloplasmin was estimated spectrophotometrically by using p-phenylenediamine (PPD) oxidase activity (Sunderman and Nomoto, 1970). Ceruloplasmin oxidizes PPD at pH 5.4 and yield a colored product, whose rate of formation is therefore, proportional to the concentration of serum ceruloplasmin.

3.2.2.7.1. Materials

- Sodium acetate solution, (0.2 mol/L).
- Acetic acid solution, (0.2 mol/L).
- Acetate buffer solution,(0.1 mol/L, pH 5.45 at 37 °C).
- Sodium azide solution, (1.5 mol/L).
- Buffered p-phenylenediamine dihydrochloride (PPD) solution (27.6’ mmol/L).

3.2.2.7.2. Methodology

- 2 ml of acetate buffer solutionwas taken into two test tubes, labelled R (reaction) and B (blank).
- 0.1 ml Serum, was added to each tube.
- Tubes R and B were placed at 37 °C in a water bath in order to reach thermal equilibrium.
- A flask containing buffered PPD solution was also placed in the water bath.
- 1 ml of warmed, buffered PPD solution was added to both the tubes. The contents of the tubes are mixed and kept unstoppered in the water bath. Care should be taken to cover the water bath for avoiding exposure of the tubes to light.
- After 5 min, 50 µl of sodium azide solution was pipetted into tube B, and the contents were mixed. The tube was replaced in the water bath.
- Exactly after 30 min, 50 µl of sodium azide solution was also added to tube R, and the contents were mixed.
- The contents of both the tubes R and B were transferred to spectrophotometer cuvets and absorbance was measured at 530 nm with a spectrophotometer. The color of the samples has stability for minimum 6 hours.

3.2.2.8. Serum Creatinine Assay

This assay is based on the direct relationship between serum creatinine and the change in absorbance over a fixed time interval, when creatinine is reacted with alkaline picrate (Lustgarten and Wenk, 1972).

3.2.2.8.1. Materials

- Working reagent prepared by mixing 0.5 mol/L NaOH and saturated aqueous picric acid (prepared at room temperature) in equal volumes.
- Creatinine standards were prepared by dissolving National Bureau of Standards creatinine (SRM No. 914) in dilute HC1 (20 mmol/L).
- In case of the reference method, the calibration curve was prepared using creatinine standards.

3.2.2.8.2. Methodology

- 100 µl samples was added to 2.0 ml of working reagent, mixed immediately and a stopwatch was started.
- At 20 s, the absorbance (A_0) of the clear solution was read and recorded.
- At 80 s a second reading (A_t) was recorded.
- The change in absorbance (ΔA) was obtained by A_0 from A_t.
- The reaction temperature was held constant at 30°C absorbance was measured at 515 nm.
- A standard curve was prepared reference information

The unknown concentration of Creatinine in a serum sample was calculated using the following formula:

$$\Delta \mathbf{A} = A_0 - A_t$$

$$\textbf{Unknown Concentration} = \frac{\Delta A\ (\text{Unknown})}{\Delta A\ (\text{Standard})} \times \text{Concentration of Standard}$$

3.3. Statistical Analyses

3.3.1. Population based study

3.3.1.1. Estimation of observed phenotypic frequency, KIR locus frequency and genotypic frequency.

The **observed phenotypic frequency (OF)** of each KIR gene is expressed as the ratio of the total number of individuals carrying that gene in the population to the total population size. The formula is as follows:

$$\mathbf{OF} = \frac{\mathrm{n(No.\ of\ individuals\ having\ the\ gene)}}{\mathrm{N(Total\ individuals\ in\ the\ population)}}$$

Estimation of **KIR locus frequency (KLF)** was made by using the formula:

$$\mathbf{KLF} = 1 - \sqrt{1 - \mathrm{f}}$$

Where f = OF of a particular KIR gene in a population.

Genotypic frequency was estimated by using the following formula

$$\mathbf{Genotypic\ frequency} = \frac{\mathrm{n(No.\ of\ individuals\ having\ a\ particular\ genotype}}{\mathrm{N(Total\ individuals\ in\ the\ population)}}$$

3.3.1.2. KIR gene frequencies of the referral populations

The KIR gene frequencies of the referral populations were taken from the following publications and 'http://www.allelefrequencies.net' database (Gonzalez-Galarza, *et al.*, 2011) as follows: South Indian populations namely Kanikars, Paravars and the Mollukurumba (Rajalingam, *et al.*, 2008), Finnish, French, Guadeloupe Caribbean, Senegal African, and Reunion having Indian Ocean origin (Denis, *et al.*, 2005), North Indian (Rajalingam, *et al.*, 2002), Samoan, Cook Island,, Tokelau, Tongan (Velickovic, *et al.*, 2006), Huichol, Mestizo, Tarahumara, Purepecha (Gutierrez-Rodriguez, *et al.*, 2006), Amazonian Amerindian (Ewerton, *et al.*, 2007), Wichis and Chiriguanos (Flores, *et al.*, 2007), Northern Irish (Middleton, *et al.*, 2007), Basque population (Santin, *et al.*, 2006), Eastern Mainland Chinese (Wu, *et al.*, 2009), Chinese Han (Jiang, *et al.*, 2005), Korean (Whang, *et al.*, 2005), Japanese (Yawata, *et al.*, 2002), Warao, Bari, Yucpa

(Gendzekhadze, *et al.*, 2006), Vietnamese and Australian Aborigine (Toneva, *et al.*, 2001), American Caucasian, Hispanic, Afro-American (Du, *et al.*, 2007), Thai, British Caucasian, Palestinian (Norman, *et al.*, 2001), Australian Caucasian (Witt, *et al.*, 1999), New York Caucasian (Hsu, *et al.*, 2002), Greeks (Niokou, *et al.*, 2003), Afro-Caribbean, Pakistani, Trinidad Asian (Norman, *et al.*, 2002), Chinese, Malay and Indian in Singapore (Lee, *et al.*, 2008), Indian Parsis and Maharashtrian (Kulkarni, *et al.*, 2008), Tibetans (Zhu, *et al.*, 2010)and Iranian Arabs and Persians (Ashouri, *et al.*, 2009).

3.3.1.3. Chi-Square (χ^2) test and estimation of significant differences

The standard χ^2 test was executed using Kyplot 2.0 beta 15 software to compare frequency differences of KIR genes between our study populations and some other geographically neighbouring populations (Kulkarni, *et al.*, 2008;Lee, *et al.*, 2008;Middleton, *et al.*, 2003;Norman, *et al.*, 2002;Rajalingam, *et al.*, 2008;Rajalingam, *et al.*, 2002;Wu, *et al.*, 2009;Zhu, *et al.*, 2010). The input file was prepared in the form of a 2X2 contingency table on which the χ^2 analysis was run. The level of significance is expressed in terms of probability (p) values wherein two-tailed p values ≥ 0.05 were considered to be statistically significant. However, differences at even 1% and 0.1% were also mentioned in our study.

3.3.1.4. Distance analyses

The frequencies of KIR genes in our populations along with that of few other geographically neighbouring populations were bootstrapped for 1000 replicates followed by calculation of Nei's genetic distances (Nei, 1972) using PHYLIP software version 3.69 (Felsenstein, 1989). Nei's genetic distance was measured by the formula:

$$D = -\ln \frac{\sum_n \sum_i p_{1ni}\, p_{2ni}}{[\sum_n \sum_i p_{1ni}^2]^{1/2}\, [\sum_n \sum_i p_{2ni}^2]^{1/2}}$$

where n is summed over loci, i over alleles at the n-th locus, and where

p_{1ni}= the frequency of i-th allele of n-th locus in population 1.

These genetic distance data were then used to construct a consensus Neighbor-Joining (NJ) tree using the Neighbor subprogram of the above-mentioned software (Felsenstein, 1989).

3.3.1.5. Principle Component analyses

Principal components analysis (PCA) identifies a smaller number of uncorrelated variables using a large set of data. These variables are known as "principal components". The number of variables that are used in the PCA are either equal to or more than the number of derived components. The first principal component measures the largest possible variance followed by the succeeding components, which in turn measures the highest possible variance under the available constraint that is not related to the preceding components. Thus, PCA explain the maximum variance using the fewest possible number of principal components..

The PCA was carried out using the MINITAB software version 16, based on the frequencies of KIR genes in our study populations and other referral populations (Ashouri, *et al.*, 2009;Denis, *et al.*, 2005;Du, *et al.*, 2007;Ewerton, *et al.*, 2007;Flores, *et al.*, 2007;Gendzekhadze, *et al.*, 2006;Gonzalez-Galarza, *et al.*, 2011;Gutierrez-Rodriguez, *et al.*, 2006;Hsu, *et al.*, 2002;Jiang, *et al.*, 2005;Kulkarni, *et al.*, 2008;Lee, *et al.*, 2008;Middleton, *et al.*, 2007;Niokou, *et al.*, 2003;Norman, *et al.*, 2002;Norman, *et al.*, 2001;Rajalingam, *et al.*, 2008;Rajalingam, *et al.*, 2002;Santin, *et al.*, 2006;Toneva, *et al.*, 2001;Velickovic, *et al.*, 2006;Whang, *et al.*, 2005;Witt, *et al.*, 1999;Wu, *et al.*, 2009;Yawata, *et al.*, 2002;Zhu, *et al.*, 2010).. The first two components were selected for preparing the score plot.

3.3.1.6. Restricted Maximum Likelihood (REML) analysis

The frequency data sets of KIR genes in our study populations and other reference populations were bootstrapped for 100 replicates and then subjected to Restricted Maximum Likelihood (REML) analyses using PHYLIP, version 3.69 (Felsenstein, 1989). This was followed by the construction of a consensus tree using FigTree version 1.3.1

(http://tree.bio.ed.ac.uk) software. The REML analyses assumed the independent evolution of each locus by pure genetic drift.

3.3.1.7. Estimation of Linkage Disequilibrium (LD) and Haplogroup analyses

The classical linkage disequilibrium coefficient (D) between pairs of KIR loci, the standardized coefficient (D'), the conventional measure of linkage disequilibrium (r^2), the two locus haplotype frequencies and the respective P values between pairs of KIR gene loci for each of the studied populations were measured. The LD values are measured using Arlequin software ver. 3.5.1.2 (Excoffier and Lischer, 2010).

Based on genotypic profile of the KIR genes in the Rajbanshis, group A and B haplotype frequencies were estimated using the following formula: group-A=$2n_{AA}+n_{AB}/2n$ and group-B=$2n_{BB}+n_{AB}/2n$. Herein, n_{AA} represents the numbers of AA genotypes and so are n_{AB} and n_{BB} for AB and BB genotypes respectively, and n represents the total size of the population.

Fourteen (14) KIR haplotypes were considered in our study which included six (6) gene content haplotypes that have been identified through segregation (Khakoo and Carrington, 2006) and analysis of sequences (Pyo, *et al.*, 2010) and were commonly found in several major ethnic groups. The other eight (8) haplotypes that were included in this study were selected from IPD-KIR database based on their uniqueness in KIR gene content. The pseudogenes namely KIR2DP1 and KIR3DP1 were not included in the haplotypes since they have no functional relevance. Moreover, KIR2DL5A and B were also excluded to avoid ambiguity since we have estimated the frequencies of KIR2DL5 as a whole and not separately for KIR2DL5A and 2DL5B respectively. The frequencies of the haplotypes were estimated using HaploIHP software generously provided by Zhang et al (Yoo, *et al.*, 2007).

3.3.2. RA based study

3.3.2.1. Estimation of clinical parameters between RA patients and controls.

The mean, standard deviation of the mean, median and Interquartile ranges of the clinical parameters were estimated using MS-Excel program of the Microsoft Office Package 2007, and the values were confirmed using IBM SPSS statistics version 19.

3.3.2.2. Diagnostic Test Evaluation

Diagnostic test evaluation criteria, which include sensitivity, specificity, positive predictive value and negative predictive value, were calculated using MedCalc, version 15.6 for Windows (MedCalc Software, Ostend, Belgium). Positive and negative cases in both the patient and the control groups were tabulated in a 2X2 table as follows:

Test	Disease (n)	Control (n)	Total
Positive	a (True Positive)	c (False Positive)	a+c
Negative	b (False negative)	d (True Negative)	b+d
Total	a+b	c+d	a+b+c+d

The values for a particular test from the above mentioned table were used to calculate the above mentioned criteria using the following formulae:

$$\textbf{Sensitivity} = \frac{a}{a+b}$$

$$\textbf{Specificity} = \frac{d}{c+d}$$

$$\textbf{Positive Predictive Value} = \frac{a}{a+c}$$

$$\textbf{Negative Predictive Value} = \frac{d}{b+d}$$

3.3.2.2.1. Odds Ratio and relative Risk.

The **odds ratio (OR) and confidence interval at 95%** were measured according to Altman, 1991 (Altman, 1991). The data of the 2X2 table under section 3.3.2.2 were used to calculate the above-mentioned values using the following formulae:

$$\textbf{Odds Ratio (OR)} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a \times d}{b \times c}$$

$$\textbf{95\% CI} = \exp(\ln(OR) - 1.96 \times SE\{\ln(OR)\}) \text{ to } \exp(\ln(OR) + 1.96 \times SE\{\ln(OR)\})$$

Where zeros caused problems with computation of the odds ratio (OR), 0.5 was added to each cell (Pagano and Gauvreau, 2000).

The **Reactive Risk or risk Ratio** was calculated according to Altman 1991 (Altman, 1991) using the following formula:

$$\textbf{Relative Risk (RR)} = \frac{a/a+b}{c/c+d}$$

Where zeros cause problems with computation of OR or RR, 0.5 is added to each cell (Pagano and Gauvreau, 2000).

3.3.2.3. Regression Analysis and Graphical Representations

Multiple regression models were computed for predicting DAS28 score in case of RA patients using ESR in combination with Anti-CCP and RF titre values. The computation was performed using IBM SPSS statistics version 19.0 software. Several softwares were used to construct the graphs and plots that were used in this part of the study, which included Sigmaplot 12.5 software, IBM SPSS Statistics version 19, Minitab version 16.1.1 and Ms-Excel program of Microsoft office package 2007. The applied software for the construction of a specific graph/plot has been mentioned in the legend of the figure representing the plot in the result section.

3.3.2.4. KIR-RA association based analyses

In this part of the RA based association study KIR gene based analyses was done wherein we followed similar statistical approaches to that of the population based study (Section 3.3.1). Herein, we have estimated the gene frequencies, genotypic frequencies, Odds ratio and linkage disequilibrium among the KIR genes in the RA patients.