

Frequency Estimation of Prothrombin Allelic Polymorphisms in Indian Population

Guttula Satyavani

Department of Applied Sciences (Bioinformatics), Indian Institute of Information Technology-Allahabad, Allahabad-211012, Uttar Pradesh, India

Abstract: The study here is to estimate the frequency of Prothrombin allelic polymorphisms from randomly selected general population. Prothrombin is a blood-clotting protein, a vitamin K-dependent clotting factor. The gene has a mutation at position 20210, hence the disorder being referred to as Prothrombin mutation 20210. The PT20210 polymorphisms were identified using simple PCR and followed by Restriction Fragment Length Polymorphism. The mutation leads to an increased amount of thrombin circulating in the person's blood stream. About 1-2% of the general population is heterozygous (one copy) for the prothrombin gene mutation. There is contradicting evidence regarding the role of the prothrombin gene mutation and arterial thrombosis. Mutations in the gene for prothrombin (F2 20210A) and factor V (F5 1691A, factor V Leiden) are established risk factors for Deep Venous Thrombosis (DVT) genetic variant in the 38-untranslated region of the prothrombin (factor II) gene (G20210A). Based on the banding patterns observed in the Fig. 1 and 2 samples are interpreted as Homozygote dominant, Homozygote recessive and Heterozygous. From the selected population, concluded that all are Homozygote polymorphic. No polymorphisms which reported to involve in thrombosis are not present in coastal Andhra population. Once large scale studies are completed confirming these results in stroke risk assessment do not need to depend on this marker any more in India.

Keywords: Prothrombin, prothrombin mutation, venous thrombosis

INTRODUCTION

Prothrombin is a blood-clotting protein. Injury to a blood vessel produces a signal which triggers the conversion of Prothrombin to thrombin. Thrombin is a protein which plays a central role in provoking the assembly of other proteins to form the blood clot. Prothrombin (F2) is a vitamin K-dependent clotting factor. Prothrombin is the precursor of the serine protease thrombin, a critical enzyme in blood coagulation (Furie and Furie, 1988). It is synthesized in the liver as a single-chain glycoprotein composed of 579 amino acid residues with a molecular weight of 72,000 (Degen *et al.*, 1983; Furie and Furie, 1988). Prothrombin (factor II) is the final effector of the clotting cascade that leads to the formation of fibrin. Prothrombin is the precursor of the serin protease thrombin which plays a central role in the control of blood coagulation by exerting both procoagulant and anticoagulant effects (Renner *et al.*, 2000).

The primary structure (Downing *et al.*, 1977; Walz *et al.*, 1977), the cDNA sequence (Degen *et al.*, 1983), the genomic DNA sequence of human Prothrombin containing 14 exons in a total length of 21 kb (Degen and Davie, 1987) and chromosomal mapping to chromosome 11p11-q12 (Royle *et al.*, 1987) have been reported. The nucleotide sequences of cDNA (Degen *et al.*, 1983) and genomic DNA (Degen and Davie, 1987) for human Prothrombin have been reported and this gene has been assigned to the chromosomal region 11p11-q12 (Royle *et al.*, 1987). Its polymorphisms have

also been described (Degen and Davie, 1987; De Vetten *et al.*, 1990; McAlpine *et al.*, 1990; Iwahana *et al.*, 1991).

Association studies also suggest that the G20210A mutation (G to A substitution at nucleotide position 20210) in the prothrombin gene (*PT*) is associated with increasing in plasma prothrombin activity and along with it, increases risk for venous thromboembolism. To test directly for linkage between this *PT* variant and plasma prothrombin activity family-based study can be performed (Williams *et al.*, 1999).

After examining the prevalence of the G20210A prothrombin mutation in a population of subjects with angiographic documentation of the condition of their coronary vessels. We studied a group of patients with severe coronary atherosclerosis, with or without a documented history of MI and a control group with normal coronary arteries. By this information we can cover the following aims like assessing the association of the G20210A mutation per se with CAD and/or MI, in order to evaluate the relationship between the G20210A mutation and plasma prothrombin levels and examining whether elevated prothrombin levels themselves were associated with CAD (Poort *et al.*, 1996).

Mechanism of action of prothrombin gene mutation: Prothrombin is the precursor to thrombin in the coagulation cascade Thrombin is required in order to

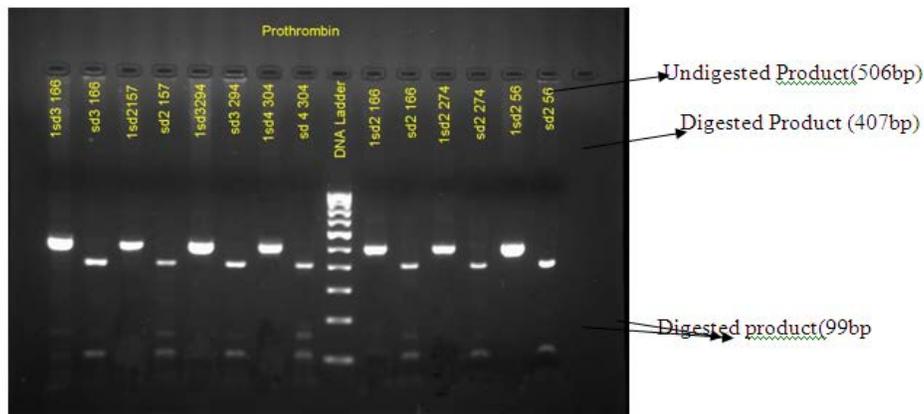


Fig. 1: Prothrombin gene amplification and RFLP by Hind III. The lanes 1, 3,5, 7, 10, 12, 14 are undigested PCR products of Prothrombin gene amplicon. 2, 4, 6, 8, 11, 13 and 15 are digested products using enzyme HindIII of the same



Fig. 2: Prothrombin gene amplification and RFLP by Hind III. The lane 1 is undigested PCR products of Prothrombin gene amplicon and 2 is digested products using enzyme HindIII of the same

convert fibrinogen into fibrin, which is the primary goal of the coagulation cascade. The gene has a mutation at position 20210, hence the disorder being referred to as Prothrombin mutation 20210. The mutation leads to an increased amount of thrombin circulating in the person's blood stream. The exact mechanism by which the Prothrombin gene mutation results in a thrombophilic state is unclear. It is thought that the increased amount of circulating prothrombin provides a springboard upon which the clotting cascade can get started and that, in some circumstances, it may run out of control because of that springboard potential (Howard *et al.*, 1997).

Epidemiology of the prothrombin gene mutation: The prothrombin gene mutation is seen more commonly in the Caucasian population. About 1-2% of the general population is heterozygous (one copy) for the prothrombin gene mutation. The Prothrombin gene mutation is relatively uncommon in the native

populations of India, Korea, Africa and North America (Rosendaal *et al.*, 1997).

Risks of the prothrombin gene mutation: The overall estimated incidence (annual occurrence) of deep venous thrombosis is 1 episode for every 1000 persons. This does not separate patients who had predisposing conditions from those who do not. At this time, there is contradicting evidence regarding the role of the prothrombin gene mutation and arterial thrombosis (stroke, heart attack). Based on these data, for persons with the prothrombin gene mutation, the most important preventive steps for the purposes of arterial disease are controlling other risk factors including: smoking, hypertension, hyperlipidemia, obesity and a sedentary lifestyle (Ferraresi *et al.*, 1997).

Diseases relation of prothrombin: Mutations in the gene for prothrombin (F2 20210A) and factor V (F5 1691A, factor V Leiden) are established risk factors for Deep Venous Thrombosis (DVT). Genetic variant in the 38-untranslated region of the prothrombin (factor II) gene (G20210A), it is apparent that this mutation confers a moderate risk for venous Thrombosis. Recently, a mutation in the gene for factor XIII (F13 100T) leading to a Valine-Leucine exchange at amino acid position 34 has been reported to be protective against DVT (Renner *et al.*, 2000).

Analysis of these mutations may help to analyze the individual risk for DVT (Ridker *et al.*, 1997).

Cardiovascular disease is the result of an interaction between environmental influences and genetic predisposition. In addition to the well-accepted traditional risk factors, there is increasing evidence suggesting that coagulation may be involved in the pathogenesis of atherosclerosis and also in the clinical progression to plaque rupture and localized occlusive thrombus formation.

The association between environmental factors and myocardial infarction has been thoroughly investigated, but the role of genetic markers is still poorly defined.

Over the past few years, studies have focused on the role of haemostatic markers, which reflect inherited or acquired propensities for thrombosis and/or the extent of subclinical atherosclerosis and several genetic mutations affecting coagulation proteins, which have been suggested as prothrombotic risk factors. Several studies have recently examined the relationship between myocardial infarction and prothrombotic genetic markers,

The Prothrombin G20210A mutation was associated with a higher prothrombin clotting activity and a 2- to 7- fold increase in risk of venous thrombosis. Other groups reported similar findings. The role of the prothrombin gene G20210A variant in CAD has not yet been established. Several investigators, however, have reported a significant increase in the prevalence of the prothrombin gene G20210A variant in patients with CAD (1.8 to 12.5%), compared with newborns or age-matched controls and a 4-fold increase in risk of myocardial infarction in young women with this variant. Others found no increased prevalence of the prothrombin gene G20210A variant in patients with CAD, compared with age and sex-matched controls.

The few studies in India showed that there are on PT20210 polymorphisms. However in Coastal Andhra There were no studies reported. In the light of above I have taken this study.

LITERATURE REVIEW

The results shown by the early study say that In general, the prevalence of the 20210A allele among Europeans ranges from 1.7 to 2% and it was found that this polymorphism is very rare (1%) in individuals of Asian and African descent. Detection of prothrombin gene polymorphism was done according to method described by Poort *et al.* (1996) Presence of 20210 G-A allele was screened by PCR followed by Hind III enzyme digestion. Forward and reverse primers used for exon 14 were 5'TCTAGAAACAGTTGCCTGGC-3'and 5'ATAGCACTGGGAGCATTGAAGC-3'. Wild type allele (20210G) lacks the restriction site and therefore remains undigested (345 bp). Positive controls for both the mutant allele were taken from recurrent spontaneous abortion (Gupta *et al.*, 2003).

A noncoding mutation in the 3'-untranslated region (UTR) of the prothrombin gene (20210G_A) has been associated with thrombophilia. Several studies have since confirmed the association of this common mutation with arterial or venous, thrombosis. The common prothrombin gene cleavage site mutation 20210G>A is associated with elevated prothrombin levels and thrombosis. The pathomechanism of the 20210G>A mutation was explained by increased mRNA formation and/or more efficient translation. Human studies also showed an influence of the intronic 19911A>G polymorphism on prothrombin activity (1). The cleavage site pattern was homogeneous with 20210A, which may cause a favourable intracellular processing and heterogeneous with 20210G. In an

independent assay for splicing efficiency, 19911G showed about 30% higher efficiency than 19911A. They conclude that the intronic 19911A>G single nucleotide polymorphism is itself functional and changes splicing efficiency by altering a known functional pentamer motif. The prothrombin G20210A mutation is thought to elevate prothrombotic risk through a gain of function, resulting in an increased level of prothrombin. Many studies have shown elevated prothrombin levels in patients with mutations, which are generally more pronounced in homozygous cases. In addition to finding elevated prothrombin levels in patients with the 20210A allele, Poort and colleagues found that the prothrombin level itself was a risk factor for thrombosis. The G20210A mutation had been shown to increase the efficiency of the 3'end cleavage signal, resulting in increased cleavage site recognition, improved processing and a more effective poly (A) site. The end result is accumulation of mRNA and increased protein synthesis and the resulting increase in prothrombin level is thought to confer the increased risk of thrombosis. The prothrombin G20210A mutation is a weak but consistent risk factor for VTEs. Homozygous G20210A mutations are rare, with 70 cases reported throughout the literature (including cases described herein). These homozygous cases display a striking phenotypic heterogeneity, from those remaining asymptomatic throughout life despite numerous other risk factors to those suffering from fatal events in the neonatal period. The current body of literature supports a multifactorial model for risk of venous thromboembolism, demonstrating that thorough clinical and laboratory evaluations must be performed and integrated to obtain an accurate picture of the overall risk for any given patient. The prevalence of the A 20210 allele of the prothrombin gene in 420 unrelated individuals (840 chromosomes) who belong to four different ethnic groups: Whites, African and Brazilian Blacks, Asians and Amerindians. PCR amplification followed by *Hind*III digestion was employed. The polymorphism was found in heterozygosity in 2 out of 120 Whites or a prevalence of 1.6% (allele frequency 0.8%), similar to that observed for other Caucasian populations. The A allele was absent among the other ethnic groups analyzed (Franco *et al.*, 1998).

The molecular basis was the increase in plasma levels of prothrombin in individuals carrying the A allele variant of the prothrombin G20210A polymorphism. Genetic studies have provided strong evidence to indicate that this particular polymorphism is indeed functional. The position of this polymorphism in the 3'-UTR led to the suggestion that this CG to CA substitution may alter either prothrombin mRNA processing, mRNA stability or translation. It has now become clear that these processes are interlinked, for example, mRNA cleavage and polyadenylation can directly and indirectly influence mRNA stability and translation initiation. Given the position of this polymorphism, it is conceptually possible that this

single nucleotide substitution influences more than one parameter of mRNA metabolism (Carter *et al.*, 2002).

Amplifying the prothrombin gene produces a 345-bp amplicon. Digestion of the 345-bp amplicon with *HindIII* restriction enzyme results in 3 different banding patterns, representing the 3 possible genotypes: G/A, A/A and G/G. Heterozygous individuals carrying the 20210 G-A prothrombin variant in one allele, genotype G/A, will generate 3 fragments of 345, 322 and 23 bp. Homozygous individuals carrying the 20210 G-A prothrombin variant in both alleles, genotype A/A, will generate 2 fragments of 322 and 23 bp. However, the lack of restriction site in normal individuals, genotype G/G (not carrying the prothrombin 20210A allele, wild type), will generate only the 345-bp fragment. Prothrombin is encoded by a 21-kilobase gene mapped to chromosome 11 (11p11-q12). Gene is organized into 14 exons separated by 13 introns and has 59 and 39-untranslated regions that are thought to play a role in its regulation and expression. A variant of the prothrombin gene (a G-A substitution at nucleotide position 20210 in the 39-untranslated region of the prothrombin gene) that was associated with elevated plasma prothrombin levels and an increased risk for venous and arterial thrombosis (Abu-Amro *et al.*, 2002). Prothrombin G20210A gene variant has been found in 0-23% of patients with portal vein thrombosis (PVT). This wide variation makes it difficult to assess the importance of prothrombin G20210A gene variant as a predisposing factor for PVT. In this study from South India, none of the patients with idiopathic PVT (0/38) or any of the controls (0/46) had prothrombin G20210A gene variant. Prothrombin G20210A gene variant does not contribute to the development of PVT in India.

Objectives: To estimate the frequency of Prothrombin allelic polymorphisms from randomly selected general population.

MATERIALS AND METHODS

The PT20210 polymorphism were identified using simple PCR and followed by Restriction Fragment Length Polymorphism according to protocol developed Lucotte and Champenois (2003). 10 mL EDTA Blood sample from 10 healthy volunteers of age group 20 to 35, DNA was isolated using salting out method. The isolated DNA was subjected to Single Polymerase Chain Reaction. Then the resultant amplicons were subjected to Restriction Fragment Analysis using *HindIII* enzyme.

The cleavage of this sequence between the AA's results in 5' overhangs on the DNA called sticky ends:



Undigested product of Prothrombin when amplified with specific primers the size is 507 bp when

amplicon is digested with *hind III* enzyme normal homozygote produces 407 and 99 bp if there is a polymorphism as G at 20210 the resulting RFLP pattern will produce 384, 99, 23 bp where as in case of heterozygote 407, 384, 99 and 23 bp bands will be resulted. In Table 1 the expected RFLP banding pattern is depicted.

RESULTS

HindIII: Is a type II site-specific deoxyribonuclease restriction enzyme isolated from *Haemophilus influenzae* that cleaves the palindromic DNA sequence AAGCTT in the presence of the cofactor Mg^{2+} via hydrolysis.

Sample collection: The peripheral blood samples were collected by VENIPUNCTURE PROCEDURE and DNA extracted by Salting Out Method and run PCR. The resultant PCR product was loaded in 2% Agarose Gel Electrophoresis the expected band size is 500 bp which is run along with the DNA ladder. From the resultant PCR product 5 uL of the sample is used for the Restriction Digestion (Table 2). Restriction Digestion of the sample using *Hind III* Restriction Enzyme The expected band size after digestion is 384 bp and 99 bp.

Based on the banding patterns observed in the Fig. 1 and 2 samples are interpreted as Homozygote dominant, Homozygote recessive and Heterozygous.

Result of Restriction digestion of population selected in Table 3.

Table 1: Prothrombin *HindIII* RFLP pattern

| Allelic type | Banding pattern |
|------------------------|--------------------------|
| Normal homozygote | 407bp, 99bp |
| Homozygote polymorphic | 384bp, 99bp, 23bp |
| Heterozygote | 407bp, 384bp, 99bp, 23bp |

Table 2: Results of Restriction digestion

| Sl No | Sample ID | Restriction digestion mix | PCR produce | Water |
|-------|-----------|---------------------------|-------------|-------|
| 1 | SD3 166r | 10 uL | 5 uL | 5 uL |
| 2 | SD2 157r | 10 uL | 5 uL | 5 uL |
| 3 | SD3 294r | 10 uL | 5 uL | 5 uL |
| 4 | SD 4304r | 10 uL | 5 uL | 5uL |
| 5 | SD2 166r | 10 uL | 5 uL | 5 uL |
| 6 | SD2 274r | 10 uL | 5 uL | 5 uL |
| 7 | SD2 56r | 10 uL | 5 uL | 5 uL |
| 8 | SD2 226r | 10 uL | 5 uL | 5 uL |

Table 3: Results of Restriction digestion

| Sl No | Sample ID | Band size detected (bp) | Results |
|-------|-----------|-------------------------|-------------------|
| 1 | SD3 166r | 407 and 99 | Normal Homozygote |
| 2 | SD2 157r | 407 and 99 | Normal Homozygote |
| 3 | SD3 294r | 407 and 99 | Normal Homozygote |
| 4 | SD 4304r | 407 and 99 | Normal Homozygote |
| 5 | SD 2 166r | 407 and 99 | Normal Homozygote |
| 6 | SD2 274r | 407 and 99 | Normal Homozygote |
| 7 | SD2 56r | 407 and 99 | Normal Homozygote |
| 8 | Sd2 226r | 407 and 99 | Normal Homozygote |

From the selected population, concluded that all are Homozygote polymorphic.

DISCUSSION

Thrombophilia is a group of conditions which are associated with an increased predisposition to thrombosis. Prothrombin G→A polymorphisms were shown in several studies and relationship with arterial disease are shown in several studies. The studies in Brazil, USA and Germany confirm association with arterial disease presence of the prothrombin 20210 polymorphism. A German group investigated 20 patients with unexplained junior stroke and found 2 patients heterozygotes for the prothrombin polymorphisms. Recently, a novel mutation in the 38-untranslated, region of the prothrombin gene (20210G to A) associated with high levels of plasma prothrombin has been reported to be a moderate risk factor for venous thrombosis, being present in heterozygous form in 5.4 to 7.3% of the patients and in 1.2 to 2.3% of the controls, with an increase in risk by threefold to fivefold.

In India Koshy and Jeyakumari (2006) showed that Prothrombin G20210A gene variant is not associated with idiopathic portal vein thrombosis in an area endemic for portal vein thrombosis Prothrombin G20210A gene variant has been found in 0-23% of patients with Portal Vein Thrombosis (PVT). In this study from South India, none of the patients with idiopathic PVT (0/38) or any of the controls (0/46) had prothrombin G20210A gene variant. Prothrombin G20210A gene variant does not contribute to the development of PVT in India. All the 10 samples analysed in our studies did not show prothrombin 20210 polymorphism as RFLP produced only 407 and 99 bp bands.

CONCLUSION

All the samples analysed are homozygote wild type. No polymorphisms which reported to involve in thrombosis are not present in coastal Andhra population. But the sample size is limited. More number of samples to be analysed for confirmed results. However, the studies from other parts of India also reported lack of this polymorphism in India. Once large scale studies are completed confirming these results in stroke risk assessment do not need to depend on this marker any more in India.

REFERENCES

Abu-Amero, K.K., C.A. Wyngaard, M. Kambouris and N. Dzimir, 2002. Prevalence of the 20210 G→A prothrombin variant and its association with coronary artery disease in a Middle Eastern Arab population. Arch. Pathol. Lab Med., 126(9): 1087-1090.

- Carter, A.M., M. Sachchithananthan, S. Stasinopoulos, F. Maurer and R.L. Medcalf, 2002. Prothrombin G20210A is a bifunctional gene polymorphism. Thromb. Haemostasis, 87(5): 846-853.
- Degen, S.J. and E.W. Davie, 1987. Nucleotide sequence of the gene for human prothrombin. Biochemistry, 26(19): 6165-6177.
- Degen, S.J., R.T. MacGillivray and E.W. Davie, 1983. Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin. Biochemistry, 22(9): 2087-2097.
- De Vetten, M., H.K. Ploos van Amstel and P.H. Reitsma, 1990. RFLP for the human prothrombin (F2) gene. Nucleic Acids Res., 18(19): 5917.
- Downing, M.R., R.J. Butkowski, J. Elion and K.G. Mann, 1977. Thrombin: Structural features related to specificity. Bibl Haematol., 44: 39-53.
- Ferraresi, P., C. Legnani, E. Quaglio, E. Castoldi, G. Marchetti, G. Palareti and F. Bernardi, 1997. Study of a G/A variation in the 38 untranslated region of prothrombin mRNA in Italian patients with venous thrombosis. Thromb. Haemostasis, 77: 379.
- Franco, R.F., S.E. Santos, J. Elion, M.H. Tavella and M.A. Zago, 1998. Prevalence of the G20210A polymorphism in the 3'-untranslated region of the prothrombin gene in different human populations. Acta Haematol., 100(1): 9-12.
- Furie, B. and B.C. Furie, 1988. The molecular basis of blood coagulation. Cell, 33: 505-518.
- Gupta, N., F. Khan, M. Tripathi, V.P. Singh, S. Tewari, V. Ramesh, N. Sinha and S. Agrawal, 2003. Absence of factor V Leiden (G1691A) mutation, FII G20210A allele in coronary artery disease in North India. Indian J. Med. Sci., 57(12): 535-542.
- Howard, T.E., M. Marusa, C. Channell and A. Duncan, 1997. A patient homozygous for a mutation in the prothrombin gene 38-untranslated region associated with massive thrombosis. Blood Coagul. Fibrin., 8: 316.
- Iwahana, H., K. Yoshimoto and M. Itakura, 1991. NcoI RFLP in the human prothrombin (F2) gene. Nucleic Acids Res., 19(15): 4309.
- Koshy, A. and M. Jeyakumari, 2006. Prothrombin G20210A gene variant is not associated with idiopathic portal vein thrombosis in an area endemic for portal vein thrombosis. Ann. Hematol., 85(2): 126-128.
- Lucotte, G. and T. Champenois, 2003. Duplex PCR-RFLP for simultaneous detection of factor V Leiden and prothrombin G20210A. Mol. Cell. Probes, 17(5): 267-269.
- McAlpine, P.J., M. Dixon, E.R. Guinto and MacGillivray, 1990. RTA: A PstI polymorphism in the human coagulation factor V (F5) gene. Nucleic Acids Res., 18: 7471.

- Poort, R.S., F.R. Rosendaal, P.H. Reitsma and R.M. Bertina, 1996. A common genetic variation in the 38-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood*, 88: 3699.
- Renner, W., H. Köppel, C. Hoffmann, K. Schallmoser, O. Stanger, H. Topla, T.C. Wascher and E. Pilger, 2000. Prothrombin G20210A, factor V Leiden and factor XIII Val34Leu: Common mutations of blood coagulation factors and deep vein thrombosis in Austria. *Thromb. Res.*, 99(1): 35.
- Ridker, P.M., J.P. Miletich, C.H. Hennekens and J.E. Buring, 1997. Ethnic distribution of factor V Leiden in 4047 men and women. *JAMA*, 277: 1305.
- Rosendaal, F.R., D.S. Siscovick, S.M. Schwartz, B.M. Psaty, T.E. Raghunathan and H.L.A. Vos, 1997. A common prothrombin variant (20210 G to A) increases the risk of myocardial infarction in young women. *Blood*, 90: 1747.
- Royle, N.J., D.M. Irwin, M.L. Koschinsky, R.T. MacGillivray and J.L. Hamerton, 1987. Human genes encoding prothrombin and ceruloplasmin map to 11p11-q12 and 3q21-24, respectively. *Somat. Cell Molec. Gen.*, 13(3): 285-292.
- Walz, D.A., V.Y. Wu, R. de Lamo, H. Dene and L.E. McCoy, 1977. Primary structure of human platelet factor 4. *Thromb. Res.*, 11(6): 893-8.
- Williams, J.T., P. Van Eerdewegh, L. Almasy and J. Blangero, 1999. Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. I. Likelihood formulation and simulation results. *Am. J. Hum. Genet.*, 65: 1134.