A Research Proposal

on

TO EXPLORE ASSOCIATION OF TGFBR1, TGFBR2, SMAD4 AND SMAD6 GENES WITH CHRONIC MYELOID LEUKEMIA

Submitted to

LOVELY PROFESSIONAL UNIVERSITY
In partial fulfillment of the requirements for the award of degree of

DOCTOR OF PHILOSOPHY (Ph.D.) IN BIOTECHNOLOGY

Submitted by: Yogender Shokeen

Supervised by: Dr Neeta Raj Sharma

FACULTY OF TECHNOLOGY AND SCIENCES
LOVELY PROFESSIONAL UNIVERSITY
PUNJAB
CONTENTS

1) Aim
2) Introduction
3) Literature Review
   - Chronic myeloid leukemia
   - Single nucleotide polymorphisms
   - TGF-β-Smad pathway
   - Imatinib Mesylate
4) Objectives
5) Materials and methods
   - Sample size
   - Inclusion criteria
   - Exclusion criteria
   - Sample collection and requirements
   - DNA extraction
   - High resolution melt (HRM)
   - SNaPshot technique
   - Expression analysis
   - Statistical consideration
6) Ethical consideration
7) References
RESEARCH PROPOSAL

AIM:

To explore the association of the TGFBR1, TGFBR2, SMAD4 and SMAD6 genes with Chronic Myeloid Leukemia

INTRODUCTION:

Chronic myeloid leukemia (CML) is a clonal myelo-proliferative disorder, characterized by enhanced proliferative capacity and prolonged survival of hematopoietic stem cells (HSCs), reduced apoptosis, and altered cell adhesion properties (1). Its incidence in US is 1.5 per 100,000, but in India its incidence rate is 0.8-2.2 and 0.6 to 1.6 in men and women respectively (2&3). Formation of the BCR/ABL oncogene, constitutes the codes for fusion tyrosine kinase (FTK) on the Philadelphia chromosome (Ph) (1). The standard treatment for CML is Imatinib mesylate, which is a tyrosine kinase inhibitor (4). However, emergence of resistance to this drug is a major problem (5).

The TGF-β signaling pathway is an essential regulator of cellular proliferation, differentiation, apoptosis, extracellular matrix remodeling of the cell, angiogenesis and inflammation. It can also modulate cell adhesion and migration of various cell types and favor the production of extracellular matrix proteins (6). The main components of this pathway are TGF-β1 protein, Smad 2 and Smad 3 (R-Smad), Smad 4 (Co-Smad), and Smad 6 and Smad 7(I-Smad) (6).

TGF-β signaling pathway is known to play an important role in many types of cancers. Boris Pache, in his article (7) shows that this pathway plays various roles in the process of malignant progression. There is growing evidence that in the later stages of cancer development, TGF-β contribute to cell growth, invasion, and metastasis and decreases host tumor immune responses (7). Subtle alteration of TGF-β signaling may also contribute to the development of cancer. TGF-β is known to effect development and/or proliferation of cancerous cells in tumors, genitor-urinary cancers and gynecological malignancies (7)

During hematopoiesis, when TGF-β signaling pathway acts in a regular manner, it is potent negative regulator of proliferation while stimulating differentiation and apoptosis. It also
plays important role in different types of hematological malignancies (8). In hematologic malignancies, resistance to these homeostatic effects of TGF-β develops. But the elevated levels of TGF-β can promote myelofibrosis and the pathogenesis of some hematologic malignancies through their effects on the stroma and immune system (8).

In CML, the BCR-Abl proteins target this pathway indirectly by targeting protein kinase Akt and transcription factor Fox O3 to impair the cytostatic effect of TGF-β1. (8). However it has also been reported that BCR-Abl expressing cells are more susceptible to TGF-β1/SMAD induced growth inhibition. At present, little is known about the how TGF-β1 and downstream SMAD transcription factors influence CML cell proliferation in the context of BCR-Abl expression (8).

Smad4 binds with combination of Smad2 and Smad3 and forms heterooligomer. It acts as co-SMAD in TGF-β signaling pathway. Smad 4 mutations are known to play important role in different types of cancer (9&10). In hematological malignancies, it has been shown that absence of Smad 4 is involved in the AML (11).

There are two inhibitory Smads in the pathway, Smad6 and Smad7. Smad6 prevents the phosphorylation of R-Smads as inhibitory action on the TGF-β-Smad pathway (6). Single nucleotide polymorphisms (SNPs) and expression levels of this gene are found to be associated with susceptibility and prognosis of various cancers (12 – 17).

TGFBR1 and TGFBR2 are the two receptors on the cell surface. They are transmembranous structure. They acts as binding site for the TGFβ1 ligand on the outer surface of cell membrane and phosphorylates the R-Smads inside the membrane (6). Mutations and differential expression in TGFBR1 and TGFBR2 are linked to many types of cancers. SNPs in these two genes are associated with susceptibility of breast cancer, colorectal cancer, gastric cancer (18-20). TGFBR1 gene is significantly more frequent in patients with RCC and transitional cell carcinoma of bladder (21). In CML, three microRNAs which are targeted to TGFBR2, show differential expression as compared to the control population (22).

Very little is known about the relation of TGF-β-Smad and BCR-ABL. No directs interaction has been reported yet between the pathway and BCR-ABL, but as mentioned it plays a very
important role in disease progression of CML and other hematological malignancies. The lack of direct interaction and still being important factor in disease makes it more interesting for the research purpose. Moreover no study has been reported yet which relates the information about the role of SNPs in these genes with susceptibility or prognosis in CML. We have decided to study genes responsible for important parts of this pathway to gather more knowledge about the relation of this pathway with the hallmark gene of CML. As mentioned SNPs in these genes have been linked to other disease including different types of cancer, but no study have been reported SNPs in these genes in CML. We will try to draw relation between SNPs in these genes, expression of these genes with reference to susceptibility of disease and response to the Imatinib mesylate.

In this study, we propose to sequence exons of the above mentioned four genes (SMAD4, SMAD6, TGFBR1 and TGFBR2) to investigate the role of TGF-β pathway in CML. We will correlate the genetic polymorphisms and expression levels of the key genes of this pathway with Imatinib response.

LITERATURE REVIEW

Chronic Myeloid Leukemia

Chronic myeloid leukemia (CML) is a clonal myeloproliferative disorder, characterized by enhanced proliferative capacity and prolonged survival of hematopoietic stem cells (HSCs), reduced apoptosis, and altered cell adhesion properties. The causative event in the leukemogenesis of CML is formation of the **BCR/ABL** oncogene, which codes for a constitutively active bcr/abl fusion tyrosine kinase (FTK) on the Philadelphia chromosome (Ph). Although bcr/abl FTK is a key molecular marker of CML, it still remains to be understood which molecular or cellular events drive translocation of **BCR/ABL** gene or initiate leukemogenesis of CML. Previous observations showing the intriguing detection of **BCR/ABL** fusion gene at a very low level in the blood of healthy people suggest that only a minor fraction of spontaneous Ph translocations progress to CML or that **BCR/ABL** fusion gene is essential but not sufficient to lead to the leukemogenesis without additional second cellular or molecular events (1).
TGF-β-Smad pathway

Transforming growth factor beta-1 (TGF-β1) is the prototypic member of a large family of structurally related pleiotropic secreted cytokines. The TGF-β signaling pathway is an essential regulator of cellular proliferation, differentiation, apoptosis, extracellular matrix remodeling of the cell, angiogenesis and inflammation (6). It can also modulate cell adhesion and migration of various cell types and favor the production of extracellular matrix proteins. The TGF-β super family consists of more than 30 related members in mammals, including 3 kinds of TGF-β, 4 kinds of activins and over 20 kinds of bone morphogenetic proteins. Subversion of TGF-β family signaling has been implicated in various human disease including autoimmune disease, vascular disorders and cancer (6).

As the TGF-β1/SMAD signaling pathway usually participates in a wide range of cellular processes and is involved in both homeostasis and oncogenesis, it acts as a negative autocrine growth factor, deregulation of TGF-β1 activation and its signaling will probably induce growth inhibition and apoptosis of tumor or leukemia cells (6).

Main components of TGF-β-Smad pathway are TGF-β1 protein, which is a ligand, TGFBR1 and TFDBR2, are the two ligand receptors, Smad 2 and Smad3 are the receptor regulated Smad (R-Smad), Smad4 is the common mediator Smad (co-Smad). Smad 6 and Smad 7 acts as inhibitory Smad (i-Smad) (6). As shown in the figure 1, the binding of extracellular TGF-β1 protein to the TGF-β type II receptor on the cell surface initiates the signaling pathway and causes the recruitment and dimerization of type II receptors. As a serine/threonine receptor kinase, the type II receptor recruits and catalyses the phosphorylation of type I receptor in a region rich in glycine and serine/threonine residues (termed the GS domain), which leads to the formation of a hetero-tetrameric complex with the ligand. The type II receptor phosphorylates serine residues of the type I receptor, whose GS domain consists of a series of about 30 serineglycine repeats. Smad 2 and Smad 3 are recruited and phosphorylated by the activated TGF receptor kinase upon, ligand binding and participate in the TGF-β1/Smad signaling pathway. Smad 4, which forms complexes with phosphorylated R-Smad and translocate into the nucleus where they act as transcription factors. I-Smad have suppressive effects on the TGF-β1-Smad signaling pathway by interrupting the actions of R-Smad and Co-Smad (6).
It has been shown that this pathway plays various roles in the process of malignant progression. It is potent inhibitor of normal stromal, hematopoietic and epithelial cell growth. There is growing evidence that in the later stages of cancer development, TGF-β is actively secreted by tumor cells and not merely acts as a bystander but rather contribute to cell growth, invasion, and metastasis and decreases host tumor immune responses (7). Subtle alteration of TGF-β signaling may also contribute to the development of cancer. TGF-β is known to effect development and/or proliferation of cancerous cells in tumors like brain tumor, head and neck tumors, lung cancer, breast cancer; gastrointestinal malignancies like biliary cancer and colon cancer (7). TGF-β pathway also have established role in several other types of cancers like gastric cancer, esophageal cancer liver cancer, pancreatic cancer, skin cancer (7). In genitor-urinary cancers, it plays a significant role in bladder cancer, prostate cancer and renal cancer (7). In gynecological malignancies, cancer of endometrium and cervix and ovarian cancer are affected by TGF-β signaling pathway (7).
During hematopoiesis, when TGF-β signaling pathway acts in a regular manner, it is potent negative regulator of proliferation while stimulating differentiation and apoptosis. It also plays important role hematological malignancies like leukemias, lymphomas and multiple myeloma (8). In hematologic malignancies, resistance to these homeostatic effects of TGF-β develops. Mutation and deletion of main oncoproteins are the main mechanism for resistance. These alterations define the tumor suppressive role for the TGF-β pathway in human hematological malignancies. But the elevated levels of TGF-β can promote myelofibrosis and the pathogenesis of some hematologic malignancies through their effects on the stroma and immune system (8).

In CML, instead of interfering with principal signaling transmitter SMAD, the BCR-Abl proteins target this pathway indirectly by targeting protein kinase Akt and transcription factor Fox O3 to impair the cytostatic effect of TGF-β1 and make BCR-Abl expressing cells resistant to TGF-β1 induced growth inhibition and apoptosis (6). However it has also been reported that BCR-Abl expressing cells are more susceptible to TGF-β1/SMAD induced growth inhibition, possibly resulting from increased SMAD promoter activity enhanced by ectopic BCR-Abl expression. The role of TGF-β1 and downstream SMAD transcription factors in CML cell proliferation is not well established (6).

SMAD4 gene is located on chromosome 18 (9) and has 12 exons (23). It acts as co-SMAD in TGF-β signaling pathway, and forms complex with SMAD2 and SMAD3 and helps them to translocate to nucleus. The majority of mutations in this gene are missense, nonsense, and frameshift mutations at the mad homology 2 region (MH2), which interfere with homooligomer formation of Smad4 protein and the hetero-oligomer formation between Smad4 and Smad2 proteins, resulting in disruption of TGF-β signaling (9). Smad 4 mutations are known to play important role in different types of cancer like pancreatic carcinoma, gastro-intestinal tumors, colorectal carcinoma, esophageal carcinoma, ovarian carcinoma, head and neck carcinoma, hepatocellular carcinoma and lung carcinoma (9&10). In hematological malignancies, it has been shown that absence of Smad 4 is involved in the AML. In this study it was shown that Smad 4 protein was absent in these cases due to the proteolytic activity directed towards the Smad4 (11).
Smad6 is located on chromosome 15. The gene has 4 exons (24). Smad6 acts as inhibitory action on the TGF-β-Smad pathway. It prevents the activation of R-Smad by phosphorylation and/or interfering with its nuclear translocation (6). Single nucleotide polymorphisms (SNPs) in this gene are found to be associated with ovarian cancer risk (12). Smad6 mRNA expression was associated with survival in Oral squamous cell carcinoma patients. Smad6 positively and Smad2 negatively predicts the prognosis for oral SCC patients and indicates a poor prognosis (13). The mRNA levels of Smad4 and Smad6, Smad4 and Smad7, and Smad6 and Smad7 in tumor samples showed a significant correlation in pancreatic ductal adenocarcinoma (14). Smad6 also has established role in other types of cancers like esophageal squamous cell carcinoma (15), hepatocellular carcinoma (16), and lung cancer (17) as well.

TGFBR1 and TGFBR2 genes are located on chromosome 9 (25) and chromosome 3 respectively (26). TGFBR1 has 9 exons (25) and TGFBR2 has 8 exons (26). Mutations and differential expression in these two important genes are linked to many types of cancers. SNPs in these two genes are associated with susceptibility of breast cancer (18). A hypomorph variant of TGFBR1 gene TGFBR1*6A is a potent modifier of colorectal cancer risk in positive manner (19). In gastric cancer, it was found that variant alleles of the promoter polymorphisms, TGFBR1 C-509I and TGFBR2 G-875A were associated with a significantly decreased risk of gastric cancer (21). TGFBR1 gene is significantly more frequent in patients with renal cell carcinoma and transitional cell carcinoma of bladder, suggesting it as a risk factor for the development of both the cancers (21). In a microRNA based study in CML, three microRNAs which are targeted to TGFBR2, show differential expression as compared to the control population. Two of them are found up-regulated in blast crisis while third one is found down-regulated, which clearly indicated its important role in disease progression in CML (22).

**Single Nucleotide Polymorphisms (SNPs)**

Single nucleotide polymorphisms, frequently called SNPs (pronounced “snips”), represent a difference in a single DNA building block, called a nucleotide. SNPs are helpful to predict an individual’s response to certain drugs, susceptibility to environmental factors
such as toxins, and risk of developing particular diseases. SNPs can also be used to track the inheritance of disease genes within families (27).

SNPs have been studied widely in relation to human health and different disease. In Colon and Rectal, SNPs in Smad2, Smad3, Smad4 TGFβ1, TGFβR1 are observed in reference to susceptibility to the disease (28). Polymorphisms are identified in STAT3 in Cervical Cancer in chinese women (29). WNT antagonist gene polymorphisms are associated in Renal Cancer (30). In Lung Cancer (NSCLC) also significant association was found between SNP in BMP2 with occurrence of disease (31). SNPs in TGFβ1 and TGFβR1 genes are studied for risk of Breast Cancer (32). Polymorphisms are identified in MDR1 gene in Acute leukemia patients among Turkish Population and related to multidrug resistance (33). SNP in FLT3 gene is associated with higher risk of myeloid leukemia (34)

In CML also several studies have been performed on SNPs in different genes. Genome-wide analysis revealed mutation of ASXL1 gene to be new molecular factor of disease progression in CML (35). Reduced risk of CML has been reported in the patients with SNP in cytochrome P-450 gene (36). Outcome of CML patients, treated with Imatinib, have been correlated with SNP in HOCT1 gene (37). Polymorphisms in HLA gene have been also associated with susceptibility of CML in different populations (38 -39). However SNPs in TGF- β pathway have not been evaluated so far in CML.

**Imatinib Mesylate**

![Molecular structure of Imatinib](image)

*FIG 2: Molecular structure of Imatinib (4).*

IUPAC name of Imatinib mesylate is 4-[[4-methylpiperazin-1-yl)methyl]-N-[4-methyl-3-[(4-pyridin-3-ylpyrimidin-2-yl)amino]phenyl]benzamide (4). The chemical structure of Imatinib mesylate is shown in fig 2. Imatinib was designed to treat Chronic Myeloid Leukemia
Imatinib is a 2-phenylaminopyrimidine derivative that functions as a specific inhibitor of a number of tyrosine kinase enzymes. It occupies the TK active site, leading to a decrease in activity (4). Imatinib mesylate binds to the amino acids of the BCR/ABL tyrosine kinase ATP binding site and stabilizes the inactive, non-ATP-binding form of BCR/ABL, thereby preventing tyrosine autophosphorylation and, in turn, phosphorylation of its substrates. This process ultimately results in “switching-off” the downstream signaling pathways that promote leukemogenesis (40).

From the collective analysis of these studies, imatinib mesylate appears to effectively induce high CHR and cytogenetic response rates (40).

During disease progression around 15% to 25% show resistance to Imatinib mesylate during different stages of disease. The reason for resistance is a number or genetic alterations acquired by CML progenitor cells due to decreased genomic stability (40).

Being the major drug for CML, Imatinib Mesylate is also used in several other diseases. Gastrointestinal stromal tumor (GIST) is another disease in which Imatinib is given as first line therapy. Ph+ Acute lymphoblastic leukemia (ALL), mastocytosis, chronic eosinophilic leukemia, chronic myelomonocytic leukemia, and a typical chronic myeloproliferative disorder are other types of cancer in which Imatinib is being prescribed. Imatinib is under trial for use in some other malignancies. In a case of chronic neutrophilic leukemia a durable and complete clinical and cytogenetic remission was induced by imatinib. Imatinib in combination with hydroxyurea shows promise as therapy for grade IV progressive glioblastoma multiforme (GBM) (41). Imatinib is also emerging as a novel drug for fibrotic disorders in pre-clinical and clinical trials. It may be available as a treatment option for fibrotic disorders in the near future (42).

RATIONALE

TGF-β-Smad pathway plays very important role in different solid tumors and hematological malignancies including CML. Despite lack of direct communication between BCR-ABL and TGF-β-Smad pathway, this pathway is found significantly related to disease progression in CML cell lines. Research on SNPs in key genes of this pathway is motivating. This will be a step towards understanding the gap of knowledge about TGF-β-Smad pathway in CML. Expected outcomes of our study will provide immense information about SNPs in
different genes of TGF-β-Smad pathway in relation to CML susceptibility. This will help us to identify novel SNPs/biomarkers for CML. It will help in determining that whether these SNPs also have role in susceptibility of CML and response to Imatinib and hence will help in understanding the role of TGF-β-Smad pathway in pathogenesis and prognosis of disease.

**OBJECTIVES:**

1) To discover genetic variations in TGFBR1, TGFBR2, Smad4 and Smad6 to examine association of these genes with Chronic Myeloid Leukemia.

2) To evaluate the role of genetic polymorphisms in TGFBR1, TGFBR2, Smad4 and Smad6 with reference to Imatinib response in Chronic Myeloid Leukemia.

3) To examine the differential expression of key proteins of TGF-β-Smad pathway in reference to Imatinib response in Chronic Myeloid Leukemia.

**METHODOLOGY AND MATERIALS**

**Sample Size**

In the beginning samples from 60 patients and 60 controls will be collected. The sample size will be extended to 160 patients and 160 controls after identification of important SNPs in the exons of the TGFBR1, TGFBR2, Smad4 and Smad6. The expression level of these genes will also be studied in the same number of patients and controls.

**Inclusion Criteria:**

1) All subjects should be at least 18 years of age (as below 18 years person are considered in pediatric and its criteria is totally different). Both male and female (43).

2) The control group will comprise of subjects who should not possess any type of tumor/cancer and also should not exhibit any genetic disorder or visible disease.

3) Patient should have confirmed BCR-ABL gene or Ph+ CML by Cytopathology or Cytogenetics (using diagnostic techniques likes FISH or PCR) (43).

**Exclusion criteria:**

1) The subject below 18 years of age. Both male and female.

2) Patients with any other pre-existing severe or unstable medical condition.
Sample Collection and Requirements

Both patients and the healthy control subjects will be required to give 5 ml of peripheral blood. The blood sample will be collected by qualified medical personnel in EDTA vials. Patients’ data about family history, his/her profession, period of working, confirmation date of disease, date of start of anti-cancer medication, date of relapse (if occurred), date of change of medication will be obtained through the questionnaires.

Personal information about study subjects will be kept confidential and managed according to the statutory requirements.

DNA Extraction

DNA will be extracted from peripheral blood sample. It will be done by manual salting out methods or standard protocol of commercially available kit. The exons/intron-exon boundaries and promoter regions of genes will be amplified from patient and control samples. The amplicons will be analyzed for any variation in the gene by using high resolution melt curves. The variation will be validated by sequencing or SNaPshot assay. Another technique can also be considered if found more suitable, cheaper and accurate for a particular SNP.

High Resolution Melt (HRM)

High Resolution Melting (HRM) is a novel, homogeneous, close-tube, post-PCR method, enabling genomic researchers to analyze genetic variations (SNPs, mutations, methylations) in PCR amplicons. It goes beyond the power of classical melting curve analysis by allowing to study the thermal denaturation of a double-stranded DNA in much more detail and with much higher information yield than ever before. HRM characterizes nucleic acid samples based on their disassociation (melting) behavior. Samples can be discriminated according to their sequence, length, GC content or strand complementarity. Even single base changes such as SNPs (single nucleotide polymorphisms) can be readily identified (44).

The most important High Resolution Melting application is gene scanning - the search for the presence of unknown variations in PCR amplicons prior to or as an alternative to sequencing. Mutations in PCR products are detectable by High Resolution Melting because they change
the shape of DNA melting curves. A combination of new-generation DNA dyes, high-end instrumentation and sophisticated analysis software allows to detect these changes and to derive information about the underlying sequence constellation (44).

**SNaPshot Technique**

SNaPshot combines the technologies of sequencing, dideoxy nucleoside triphosphates (ddNTP) marked with fluorescence and PCR and primer extension to determine each allele. The region surrounding the SNP was amplified by multiplex PCR. A single base extension (SBE) reaction is performed in the system including one sequencing enzyme, PCR production modals and different oligos near the SNP. The small different DNA products generating in this post-PCR primer extension reaction can be distinguished by ABI sequencer by the color of the peak, then the gene type of sample and the relative SNP site can be identified by the position of peak in gel (45). SNaPshot, a genotyping method is used to screen and confirm SNPs and assess DNA methylation (45).

**Expression Analysis**

Expression profiling will be carried out for all the genes by reverse transcription quantitative real time PCR (RT-PCR) to examine the differential expression in CML patients as compared to healthy individuals. A good quality RNA is the pre-requisite for RT-PCR experiments. RNA will be isolated from peripheral blood using TRIzol protocol. The purity of RNA will be checked by gel electrophoresis and nanodrop spectrophotometer. The RNA template will be copied to cDNA using first strand cDNA kit. qPCR will be performed on Real-time PCR machine using SYBR GREEN.

Second approach to validate the protein expression will be western blotting. Western blotting identifies with specific antibodies proteins that have been separated from one another according to their size by gel electrophoresis. The proteins in the gel are transferred to nitrocellulose or PVDP (polyvinylidene fluoride) membrane by applying electric current. The membrane is then probed using specific primary antibody (46).

**Statistical considerations**

The significance of differences in allele frequencies and genotypes between CML cases and healthy controls will be calculated using statistical tests like generalized Fisher’s
exact test, Pearson’s χ² test and Mann-Whitney U-test. The statistical analysis will be done using commercial software packages (SPSS, SAS etc.). P-values (< 0.05) will be considered statistically significant. A qualified biostatistician will be consulted for the statistical part of this research.

ETHICS CONSIDERATION

A written informed consent will be taken from all the patients and healthy volunteers for use of their blood sample for research. Updated versions of informed consent will be voluntary signed and dated by the study subject / his or her relative before any study related procedures begins. The original signed and dated informed consent will be kept in the investigator’s study file.

Formal ethics committee approval is required and will be obtained from the Ethics Committee of Sir Ganga Ram Hospital before initiating the work.

REFERENCES:


-----------------------------------------------------------------X-----------------------------------------------------------------