<table>
<thead>
<tr>
<th>Chapter No.</th>
<th>Topic</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Review of Literature</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>2.1 Mutation</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>2.2 Human Genome: Set Of Instructions For Life</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>2.3 Proteins</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>2.3.1 Protein structure</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>2.3.2 Structural information in the analysis of mutations</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>2.4 Human genome variation</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>2.5 Single nucleotide polymorphism: the simplest mutation</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2.5.1 SNPs in Coding Region</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.1.1 Non-synonymous SNPs (nsSNPs)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.1.2 Synonymous SNPs</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.1.3 Frameshift mutations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.1.4 Nonsense mutation</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.2 SNPS in non-coding region</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.2.1 Splice site variations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.2.2 Upstream variations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.2.3 5' UTR variations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.2.4 3' UTR variations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.4.5 Intrinsic variations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.4.6 Downstream variations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.5.4.7 Intergenic variations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.6 Molecular Effects of Missense Mutations</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.6.1 Effect of mutations on the functional sites</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.6.2 Effect of mutation on sequence conservation</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.6.3 Effect of mutation on PTM sites</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2.6.4 Effect of mutations on solvent accessibility</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2.6.5 Effect of substitutions involving glycine &amp; proline residues</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2.6.6 Mutation effecting structural disorder</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>2.6.7 Effect of mutation on protein-protein, protein-DAN, protein-RNA interaction</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>2.6.8 Effect of mutation on subcellular localization and protein expression</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>2.6.9 Mutation effecting misfolding and aggregation</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>2.6.10 Mutation effecting electrostatic surface potentials</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>2.6.11 Predictors of pathogenicity</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>2.6.12 Prior studies on in silico prediction of missense mutation</td>
<td>22</td>
</tr>
<tr>
<td>2.7 The human kinome</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.7.1 Key metabolic switches</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>2.7.2 Classification of protein kinases</td>
<td>24</td>
</tr>
</tbody>
</table>
2.8 Kinase domain
  2.8.1 The sequence
  2.8.2 A common feature of all protein kinases: the PK domain
    2.8.2.1 P-loop
    2.8.2.2 The αC-helix
    2.8.2.3 αF-helix
    2.8.2.4 The catalytic loop
    2.8.2.5 The activation loop (T-loop)
    2.8.2.6 P+1 loop
    2.8.2.7 The substrate-binding groove
    2.8.2.8 The ATP-binding pocket

2.9 Mining kinase mutations from the literature
  2.9.1 dbSNP
  2.9.2 Ensembl
  2.9.3 The HapMap project
  2.9.4 OMIM Online Mendelian Inheritance in Man
  2.9.5 SwissProt Variant pages and the ModSNP database
  2.9.6 SAApdb
  2.9.7 COSMIC: The Catalogue of Somatic Mutations in Cancer
  2.9.8 KinMutBase
  2.9.9 MoKCa - Mutations of Kinases in Cancer

2.10 Protein Kinases as a potent drug targets

2.11 The mechanism of drug resistance and relapse in protein kinases

2.12 Identifying missense mutations using *in silico* based approaches
  2.12.1 Sequence-based methods
    2.12.1.1 SIFT
    2.12.1.2 PANTHER
    2.12.1.3 SNPs&GO
    2.12.1.4 PhD-SNP
  2.12.2 Structure based approach
  2.12.3 Sequence vs structure based methods
    2.12.3.1 PolyPhen 2.0
    2.12.3.2 I Mutant 3.0
    2.12.3.3 SNAP
    2.12.3.4 Align-GVGD

2.13 Mapping the significant missense mutation onto 3D structure and their effects
  2.13.1 Energy minimization
  2.13.2 Docking
    2.13.2.1 AutoDock

2.14 Molecular dynamics simulations
  2.14.1 Historical background
2.14.2 Molecular dynamics methods 48
  2.14.2.1 Simulated annealing 48
  2.14.2.2 Langevin dynamics 49
  2.14.2.3 Brownian dynamics 49
  2.14.2.4 Monte Carlo 50
  2.14.2.5 Principal component analysis 51
2.14.3 Computational aspects of MD simulation 52
  2.14.3.1 GROMACS 52
  2.14.3.2 NAMD 53
  2.14.3.3 AMBER 53
  2.14.3.5 CHARMM 53
2.15 Various Protein kinases analyzed in our study 54
  2.15.1 PK(L/R), M2 54
  2.15.2 ATM 55
  2.15.3 FGFR 55
  2.15.4 ALK 56
  2.15.5 RET 57
  2.15.6 EGFR 58
3 Materials and method 60
  3.1 Dataset 60
  3.2 Profiling method 60
    3.2.1 SNPS in coding region 60
      3.2.1.1 SIFT 61
      3.2.1.2 PolyPhen 2.0 61
      3.2.1.3 I-Mutant 3.0 61
      3.2.1.4 Align-GVGD 62
      3.2.1.5 SNAP 62
    3.2.2 SNPS in regulatory region 64
      3.2.2.1 SRSide 64
      3.2.2.2 FASTSNP 64
      3.2.2.3 UTRscan 65
  3.3 In silico robustness analysis of mutations 65
  3.4 Prediction of Post Translational Modification sites 66
  3.5 Conservation analysis 67
  3.6 Ranking SNPS based on integrative scoring system 68
  3.7 Prediction Assessment 68
  3.8 Structural Modelling and Docking Analysis 69
  3.9 Docking 70
  3.10 Molecular Dynamics Simulation 71
  3.11 Analysis of molecular dynamics trajectory 71
  3.12 Principal Component Analysis 71
4 Results and Discussion 73
  4.1 Path to Facilitate the Prediction of Functional Amino Acid Substitutions In Red Blood Cell Disorders- A Computational Approach 73
    4.1.1 Selection of SNPs for analysis 73
4.1.2 Analysis of deleterious nsSNPs using SIFT
4.1.3 Analysis of deleterious nsSNPs using Polyphen
4.1.4 Identification of functional nsSNPs using I-MUTANT 2.0
4.1.5 Prediction of deleterious nsSNPs Using PANTHER
4.1.6 Combination of the prediction programs for deleterious nsSNPs
4.1.7 Characterization of functional regulatory elements in 5’ and 3’ UTR by UTRSCAN
4.1.8 Prediction of functional nsSNPs by FASTSNP
4.1.9 Modeling of deleterious nsSNPs
4.10 Glucose 6 phosphate dehydrogenase (G6PD)
4.11 Pyruvate kinase isoforms (PKLR & PKM2)
4.12 Computation of stabilizing residues in native and mutant modeled structures of G6PD and PK genes
4.13 Discussion

4.2 Computational refinement of functional Single Nucleotide Polymorphisms associated with ATM gene
4.2.1 SNP Dataset
4.2.2 Prediction of deleterious nsSNPs in coding region
4.2.3 Analysis of functional SNPS in the regulatory region
4.2.4 Concordance analysis between SIFT and Polyphen
4.2.5 Integrative ranking system and of coding nsSNPS
4.2.6 Analysis of nsSNPS in conserved region
4.2.7 Prediction of post-translational modification sites
4.2.8 Discussion

4.3 Disease-causing mutation in extracellular and intracellular domain of FGFR1 protein: computational approach
4.3.1 Dataset
4.3.2 Mutations predicted to be deleterious
4.3.3 Sequence conservation
4.3.4 Mutations predicted to affect β-aggregation
4.3.5 Mutations predicted to affect stabilizing residues
4.3.6 Mutation predicted to affect protein stability
4.3.7 Structural analysis of FGFR1 mutation
4.3.8 Docking study
4.3.9 Molecular dynamics simulation
4.3.10 Discussion

4.4 Predicting the impact of deleterious mutations in the protein kinase domain of FGFR2 in the context of function, structure, and pathogenesis – A Bioinformatics Approach
4.4.1 SNP annotation
4.4.2 Prediction of deleterious mutations in FGFR2 domains
4.4.3 Comparison of the impact of mutations in fgfr2 kinase domain
4.4.4 In silico mutational robustness analysis in protein