**FEBRUARY:**

**----------------------------------------------------------------------------------------------------**

**week 1:**

**----------------------------------------------------------------------------------------------------**

paper 1 # Sequencing technologies - the next generation.

paper 2 # Exome sequencing as a tool for Mendelian disease gene discovery.

paper 3 # Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data.

paper 4 # A framework for variation discovery and genotyping using next-generation DNA sequencing data.

paper 5 # Initiating Human Variome Project Country Node.

### Korean Mutation Database (KMD) -- backend documentation not available.

paper 6 # dbNSFP: A lightweight Database of Human Nonsynonymous SNPs and Their Functional Predictions.

### phyloP score, SIFT score, LRT score, Polyphen2 score and MutationTaster score

**week 2:**

**----------------------------------------------------------------------------------------------------**

paper 7 # The UCSC Genome Browser database: extensions and updates 2011.

### http://genome-preview.ucsc.edu -- have the sql file for the database.

paper 8 # A New Face and New Challenges for Online Mendelian Inheritence in Man (OMIM).

### http://www.omim.org -- backend documentation not available. Need to send request for any download.

paper 9 # Ensembl 2012.

### http://www.ensembl.org/info/docs/api/index.html

### Documentation available for all.

paper 10 # Ensemble BioMarts: a hub for data retrieval across taxonomic space.

### Tables are available at ftp://ftp.ensembl.org/pub

### Queries ..not sure...need to check.

paper 11 # International Cancer Genome Consortium Data Portal - a one-stop shop for cancer genomics data.

### BioMarts.

paper 12 # Integrative Cancer Genomics (IntOGen) in BioMart.

paper 13 # The 2012 Nucleic Acids Research Database issue and the online Molecular Biology Database Collection.

paper 14 # A database and API for variation dense genotyping and resequencing data.

### http://www.ensembl.org/info/docs/Pdoc/index.html

### http://www.ensembl.org/info/docs/api/index.html

### http://cvs.sanger.ac.uk/cgi-bin/viewcvs.cgi/ensembl-variation/schema

### http://www.ensembl.org/info/docs/api/variation/index.html

### http://www.ensembl.org/info/docs/api/variation/variation\_tutorial.html

paper 15 # CanProVar: A Human Cancer Proteome Variation Database.

paper 16 # The Roche Cancer Genome Database (RCGDB).

paper 17 # ANNOVAR: functional annotation of genetic variants from high-throughput seqeuencing data.

paper 18 # Querying and computing with BioCyc databases.

### BioCyc - Frame Repres- entation System (FRS) -- not useful.

paper 19 # BioMart: a data federation framework for large collaborative projects.

### http://www.biomart.org -- checked!

SETUP # BioMart and MySql - Done!!

**week 3:**

**----------------------------------------------------------------------------------------------------**

paper 20 # BioMart Central Portal: an open database network for the biological community.

SETUP # test connection between BioMart and MySql - Done!!

paper 21 # BioMart: driving a paradigm change in biological data management.

paper 22 # The Human Gene Mutation Database: 2008 update.

paper 23 # EnsMart: A Generic System for Fast and Flexible Access to Biological Data.

### Checking the data files!!

**week 4:**

**----------------------------------------------------------------------------------------------------**

paper 24 # Predicting Deleterious Amino Substitutions.[1]

paper 25 # Accounting for Human Polymrphisms Predicted to Affect Protein Function.[2]

paper 26 # Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.[3]

### SIFT

### Documentation on SIFT - Done !!

### Initial database design !!

paper 27 # A method and server for predicting damaging missense mutations.[4]

### polyphen2

### Documentation on polyphen2 - Done !!

paper 28 # New Methods for Detecting Lineage Specific Selection

### **Documentation on PhyloP -- NEEDS TO BE DONE!!**

### Documentation on MutationTaster - Done !!

paper 29 # Identification of deleterious mutations within three human genomes.[5]

### Documentation on LRT - Done !!

**week 5:**

**----------------------------------------------------------------------------------------------------**

### OMIM database check !!

### DGV database check !!

### 1000genomes database check !!

### GVS database check !!

**MARCH:**

**----------------------------------------------------------------------------------------------------**

paper 30 # F-SNP: computationally predicted functional SNPs for disease association studies.

paper 31 # An integrative scoring system for ranking SNPs by their potential deleterious effects.

### F-SNP database.

**week 6:**

**----------------------------------------------------------------------------------------------------**

### Discussion session - 1

paper 32 # GWASdb: a database for human genetic variants identified by genome-wide association studies

### GWASdb checked!!

### database prototype 2!!

**week 7:**

**----------------------------------------------------------------------------------------------------**

### getting data from NCBI for integration!!

### <http://gene2mesh.ncibi.org/>

### getting data for MeSH from NCBI for integration!!

### getting data for GO from NCBI for integration!!

For Expression DB

###http://www.proteinatlas.org/about/download