Building an ESCC database to link clinical property with molecular mechanistic determinant

**Introduction**

In order to facilitate biomedical analysis of esophagus squamous cell carcinoma (ESCC), we are building a knowledgebase that links clinical properties including the diagnosis end points to the known molecular information, i.e. genomic variants (snv, indel, and other genomic abnormality) to render pathway and network information for biomedical research and provide basis for diagnosis and relapse prediction model

**Main implementation**

Throughout this research process, we plan to implement following parts

1. A web portal for data collection
2. Design a database schema
3. Determine the information collection protocol
4. A knowledge database to store molecular and clinical information
5. A web development for information retrieval and report, a user level web portal for mechanistic analysis (similar to DAVID, IPA, Jasper etc.)

Public available databases

Network of Cancer Gene: <http://ncg.kcl.ac.uk/index.php>

**Development staging**

Determine the information collection standard/protocol

1. Explore currently existing molecular knowledge database
2. Explore the matured consortium results: 1000 genome, the Broad Institute, TCGA

Design a database schema

Develop a web portal for data collection

Build a knowledge database

Develop a web portal for use level’s access

1. A web portal for data collection
2. Design a database schema
3. Determine the information collection protocol
4. A knowledge database to store molecular and clinical information
5. A web development for information retrieval and report, a user level web portal for mechanistic analysis (similar to DAVID, IPA, Jasper etc.)