Discovery of driver genes

One crucial challenge for understanding the molecular mechanisms underlying cancer is to distinguish driver mutations and driver genes to passenger ones. Driver genes contribute to cancer initiation and development, while passenger genes accumulate in cancer cells but do not contribute to the carcinogenesis. Previous researches indicate that driver gene set has two key properties: (i) High coverage in a large number of samples; (ii) High mutual exclusivity (i.e. a single mutation is enough to disturb one pathway). These two properties can be used for the identification of cancer driver genes.

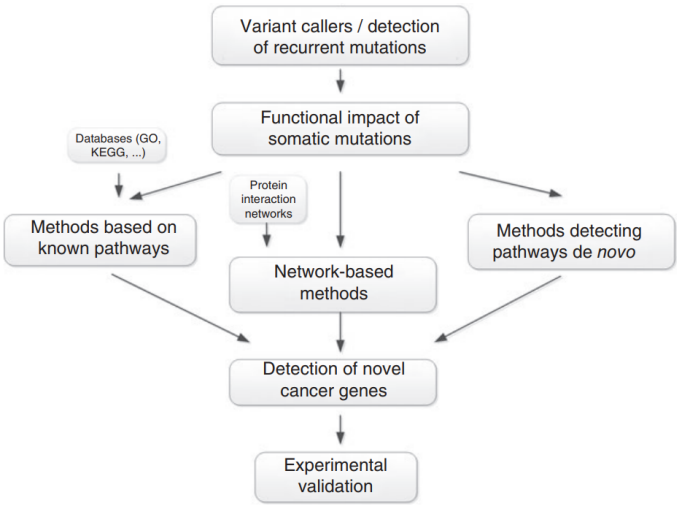


Figure 1. Detection of novel cancer genes.

Table 1. Methods for identification of driver genes in cancer.

Method	Data	Technique
MuSiC	SNV	Statistical test
MutSigCV	SNV	Statistical test
Oncodrive-CLUST	SNV	Clustering
MADGiC	Somatic mutation	Bayesian model
CADD	General variants	Support vector machine
GSEA	Mutation and known pathways	Statistical test
HotNet	Mutation and interaction networks	Heat diffusion process and statistical test
NetBox	Mutation and interaction networks	Community detection
ActiveDriver	Mutations in protein signalling sites	Statistical model
OncodriveFM	Somatic mutations	FM bias
OncodriveFML	Somatic mutations	FM bias
DeepDriver	Somatic mutation	Deep convolutional neural networks

45-1 Ramsey Road, Shirley, NY 11967, USA
www.cd-genomics.com
Tel: 1-631-275-3058 Fax: 1-631-614-7828
Email: info@cd-genomics.com

Discovery of pathways

A number of evidences have suggested that pathways usually function cooperatively in cancer initiation and progression. Therefore, exploring the collaboration among different biological pathways and functional modules probably sheds light on the molecular mechanisms underlying carcinogenesis. A number of methods have been developed for the identification of driver gene sets in cancer.

Table 2. Examples of methods for identification of driver pathways in cancer.

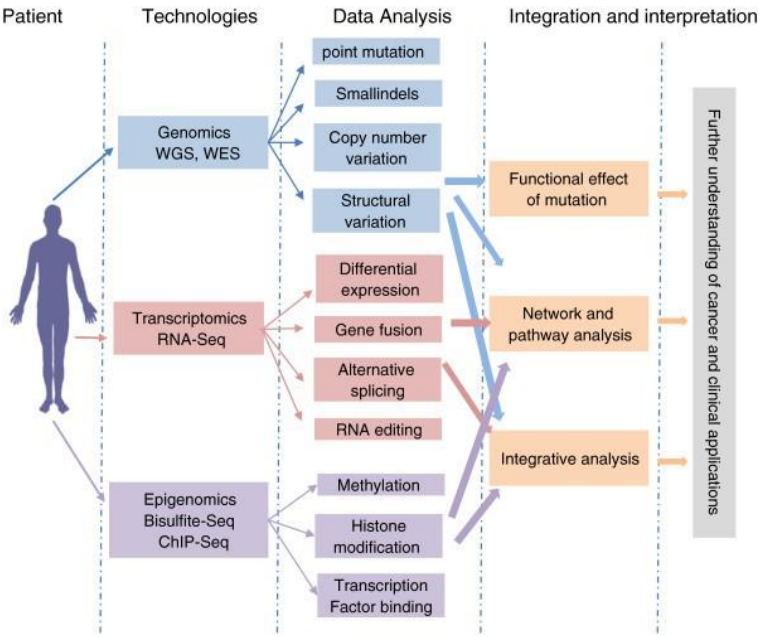
Method	Application
CoMDP	*De novo* discovers co-occurring mutated driver pathways in cancer by introducing a novel weight function.
GAMToC	Genomic Alteration Modules using Total Correlation (GAMToC) integrates copy number and mutation data to discover structural module of gene.
PSMP	Mutual exclusivity-based pairwise search for patterns of somatic mutations.
Multi-Dendrix	Simultaneously identify multiple driver gene sets in cancer based on the two properties of driver gene sets.
RME	Detect functional modules based on the patterns of recurrent and mutually exclusive aberrations.
TiMEx	A generative probabilistic model for de novo detection of mutual exclusive cancer alterations.
MEMo	Detects candidate driver subnetworks of aberrant genes with mutually exclusive patterns.
MEMCover	Uncover pan-cancer dysregulated pathways by combining across tissue type exclusivity with interaction data
HotNet2	Use insulated network diffusion to detect mutated subnetworks with statistically significant size. Captures the directionality of interactions.
ComMDP and SpeMDP	*De novo* identifies cancer common and specific driver gene sets, respectively.

45-1 Ramsey Road, Shirley, NY 11967, USA
www.cd-genomics.com
Tel: 1-631-275-3058 Fax: 1-631-614-7828
Email: info@cd-genomics.com

CD Genomics
The Genomics Services Company

NGS Approaches

NGS can detect the genomic, transcriptomic, and epigenomic changes including mutations, copy number variations (CNVs), structural variants, fusion genes, differentially expressed genes, and DNA methylation changes, *etc*.



- **Cancer Genomics --** Whole genome sequencing, whole exome sequencing, and targeted sequencing (such as gene panel sequencing) can be used to capture genomic changes in tumor tissue. Coupled with powerful bioinformatics, cancer genomics allows researchers to understand oncogenes, tumor suppressor genes, and other genetic factors that contribute to tumor initiation and progression.

- **Cancer transcriptomics --** Monitoring changes in gene expression is important to identify and characterize oncogenes. Related sequencing technologies include RNA-seq, targeted RNA sequencing, single-cell RNA sequencing, and so on. Tools such as WGCNA (weighted correlation network analysis) can be used to construct gene network, allowing to identify driver genes.

- **Cancer epigenomics --** Epigenomics technologies (such as whole genome bisulfite sequencing and ChIP-seq) can be used to explore methylation abnormalities and transcription factor binding in tumor, allowing researchers to identify which methylations are crucial for tumor initiation and development.

45-1 Ramsey Road, Shirley, NY 11967, USA
www.cd-genomics.com
Tel: 1-631-275-3058 Fax: 1-631-614-7828
Email: info@cd-genomics.com

CD Genomics
The Genomics Services Company

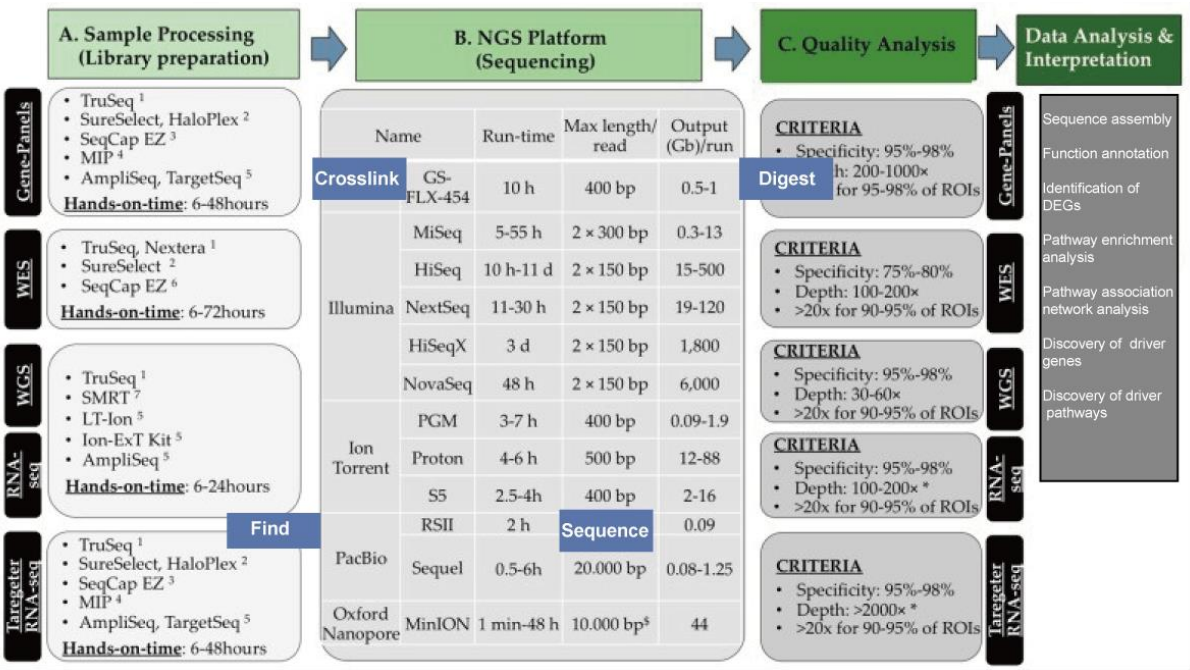
Our cancer genomics research solutions

CD Genomics offers the sequencing approaches above using next-generation sequencing or long-read sequencing technologies to explore multiple types of genetic mutations in tumor, hence understanding the molecular mechanisms underlying carcinogenesis.

I. The workflow of NGS consists of four steps: library preparation, sequencing, quality analysis, and data analysis and interpretation.

II. We have advanced sequencers from Roche, Illumina, Ion Torrent, Pacbio, and Oxford Nanopore.

III. The bioinformatics analyses probably include sequence assembly, function annotation, identification of DEGs (differentially expressed genes), pathway enrichment analysis, pathway association network analysis, discovery of driver genes and pathways, *et al*.



We can tailor this pipeline based on your specific research project.

45-1 Ramsey Road, Shirley, NY 11967, USA
www.cd-genomics.com
Tel: 1-631-275-3058 Fax: 1-631-614-7828
Email: info@cd-genomics.com