Report 1

Paper Reading

I. Impact of genetic variation on three dimensional structure and function of proteins.

1. What is the biological question?

Most studies that analyze the relationship between point mutations and 3D protein structure have been restricted to individual proteins or single diseases.

The author identified a benchmark dataset of protein structures and present a detailed overview about the observed effects of SNVs on the structure, function, stability, and binding properties of proteins.

2. What is the method?

Construction of the dataset: Based on a specific procedure, a final benchmark dataset of 374 unique human SNVs is determined, each corresponding to a different PDB entry for which 3D atomic level coordinates are available.

Manual annotation of SNVs: Systematically reviewed the available literature and search in several databases like RCSB PDB, dbSNP and PubMed. Additionally, some tools were developed to facilitate mapping of any genetic location onto corresponding protein sequences and 3D protein structures

3. What significant scientific contribution does the paper make?

The paper presents a systematic and qualitative analysis of some proteins for which 3D structures of genetic variant proteins have been determined.

A wide range of structural and functional changes caused by single amino acid differences were observed, including changes in activity, aggregation, stability, binding/dissociation, assembly, rearrangement.

These findings provide insights into atomic-level structural differences caused by the genetic variation and improve our understanding of the relationship between point mutations and experimentally observed consequences in 3D.

II. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.

1. What is the biological question?

- Due to the complexity of clinical NGS testing, tumor genomic profiling for large populations of individuals with cancer is hard to achieve.

- The nature of genomic alterations observed in individuals with advanced metastatic cancer may differ substantially from what has been characterized in primary, untreated cancers.

- The true clinical utility of mutation profiling remains uncertain, requiring careful evaluation of the degree to which molecular results influence therapeutic decisions in different clinical contexts.

This paper tried to solve the above problems through establishing a large-scale, prospective clinical sequencing initiative using a comprehensive assay, MSK-IMPACT.

2. What is the method?

MSK-IMPACT clinical workflow:
1. Patients consent 2. Sample accessioning 3. Sample preparation
4. Sequencing 5. Bioinformatic analysis 6. Case review and sign pout
7. Reported as the electronic medical record

This method allows for the detection of protein-coding mutations, copy number alterations (CNAs), and selected promoter mutations and structural rearrangements in cancer-associated genes. Moreover, A key feature of this method is the use of normal controls matched to patient tumors, helping compile a comprehensive catalog of definitively somatic (i.e., tumor-specific) mutations for every tumor sequenced.

3. What significant scientific contribution does the paper make?

Compiled tumor and matched normal sequence data from a unique cohort of more than 10,000 patients with advanced cancer and available pathological and clinical annotations.

Using these data, clinically relevant somatic mutations, novel noncoding alterations, and mutational signatures that were shared by common and rare tumor types are identified.

The full dataset is publicly accessible, facilitating discovery of novel biomarkers and deeper investigation into rare alterations and tumor types.

Protein mutation exploration

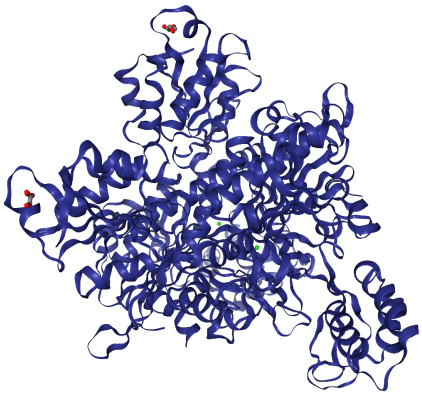
Protein of interest: human glycyl-tRNA synthetase

Mutation of interest: G526R (PDB ID: 2PMF)

Associated disease: Charcot-Marie-Tooth disease

Available information in PDB:

- There are 15 entries for this protein.
- X-ray diffraction was used in characterization of the structure.
- 3D structure (color by chain)



- Number of polypeptide chains: 1
- Sequence Length: 693
- Ligands: CHLORIDE ION, GLYCEROL

Publication record

- Paper: Wei Xie, Leslie A. Nangle, Wei Zhang, Paul Schimmel, Xiang-Lei Yang. Long-range structural effects of a Charcot–Marie–Tooth disease-causing mutation in human glycyl-tRNA synthetase. Proceedings of the National Academy of Sciences. 2007; 104(24): 9976-9981;

- Effect of this mutation: The mutation is at the site for synthesis of glycyl-adenylate, but the rest of the two structures are closely similar. Direct experiments confirm the tighter dimer interaction of the mutated protein.

Uniprot ID: P41250