Proteogenomic characterization of colorectal cancer using the IndivuType multiomics database

Margarita Krawczyk, Silvia von der Heyde, Vincent Piras*, Alexander Hergovich*, Julia Bischof, Jonathan Woodsmith, Mark Whittaker* and Hartmut Juhl
Indivumed GmbH, Falkenried 88, Bldg. E, 20251 Hamburg, Germany
* Evotec SE, Essener Bogen 7, 22419 Hamburg, Germany

Introduction

We have developed IndivuType, a knowledge and discovery platform that combines genomics, transcriptomics, and proteomics datasets to enable cutting edge precision medicine approaches.

Here we use this resource to drive an integrated proteogenomic approach to aid identification of putative therapeutic targets in a colorectal cancer (CRC) cohort of 500 patients, consisting of 388 patients with primary tumor and matching adjacent normal samples, as well as 112 patients with metastatic CRC.

We identified hundreds of proteins dysregulated in CRC that are dependent on the genetic background of the patient. Further, we characterized these proteins using clinical parameters from the IndivuType database to prioritize targets for precision medicine approaches. This project highlights the utility of moving beyond solely genomics approaches to better understand the molecular mechanisms underpinning cancer.

Genomics

Methods

- Whole genomes were sequenced to a depth of 35x (normal) and 70x (tumor/metastasis) using a NovaSeq 6000.
- Short somatic variation dataset was generated using a consensus calling approach from four variant callers.

Results

- Identification of many known and novel protein coding somatic variants (Figure 1) that can be used to define CRC sub-cohorts (Figure 2).

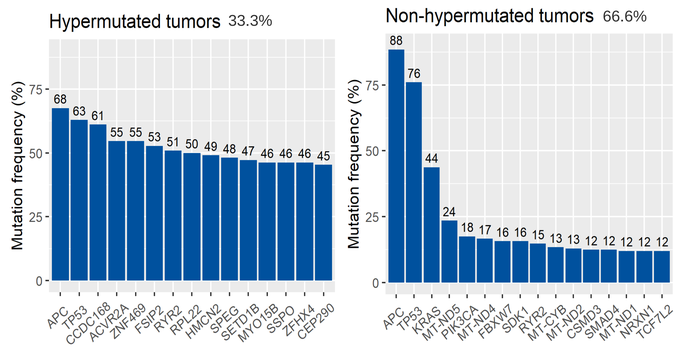

Figure 1. One third of CRC tumors in our data are hypermutated. We therefore split the cohort to reveal coding mutational patterns across distinct mutation types. The mutation frequency of all hypermutated and non-hypermutated tumors is shown. KRAS, TP53 and APC are among the most frequent ones.

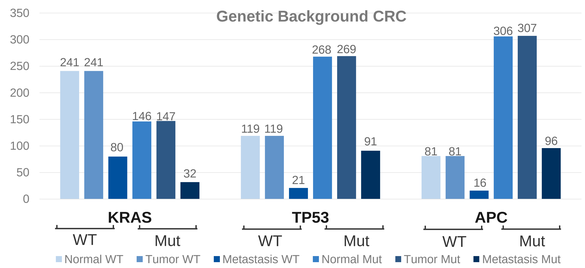

Figure 2. Distribution of WT and mutant cases for KRAS, TP53 and APC in normal, tumor and metastatic samples of 500 CRC cases.

Proteomics

Methods

- Tissue lysed and homogenised prior to Trypsin and LysC protein digestion.
- Electrospray-ionization followed by MS/MS (quadropole/orbitrap) using CID on FUSION-LUMOS.
- Data dependent acquisition of ions prior to peptide identification using MaxQuant software.
- Differentially expressed proteins identified between the three types of samples; adjacent normal, primary tumor and metastatic tumor.

Results

- Identification of hundreds of dysregulated proteins across the CRC cohort (Fig. 3A and B), encompassing many known and novel protein underlying CRC biology.
- Most significantly differentially expressed proteins are upregulated. As expected, the least differences are found in metastasis vs tumor samples.

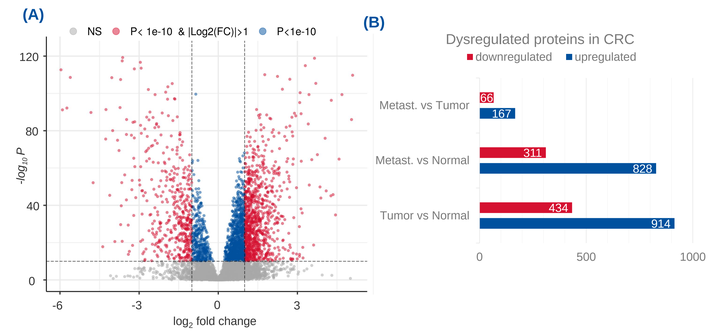

Figure 3. (A) Volcano plot of protein fold changes vs. p-value in tumor vs normal samples and (B) the corresponding quantifications, additionally with metastatic samples.

Proteogenomics – APC mutation

- Using the defined sub-cohorts of patients with a wild type and mutant APC background, we can observe the impact of the genetic background on protein expression, using normal tissue as a control.
- Of 2166 differentially expressed proteins (Fig. 4A), 1566 are significantly upregulated in tumor, but not in normal samples, of which 711 show the same trend in metastasis (Fig. 4B). Only 184 genes are downregulated of which 64 show the same trend in metastatic samples.

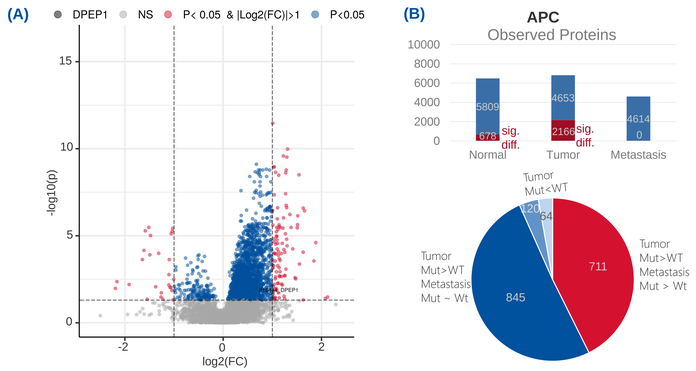

Figure 4. (A) Volcano plot of protein fold changes in APC mutated tumor samples vs non-mutated samples against p-value. (B) We observe 2166 significantly differentially expressed proteins in APC mutated vs non-mutated cases within the tumor samples, but only 678 in normal samples. We see no significant differences in protein expression in metastatic samples.

Example of APC mutant dependent expression: DPEP1

Dipeptidase 1 (DPEP1) is involved in metabolism of glutathione by dipeptide hydrolysis. DPEP1 over-expression is reported to cause significant increase in colon cancer cell adhesion, invasion and metastasis [1].

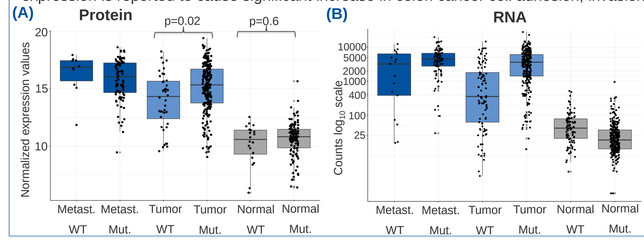

Figure 5 DPEP1 is differentially expressed in APC mutant vs WT cases in tumor samples, but not in normal samples, in (A) protein expression data and (B) RNA expression data.

Relapse-free Survival

- Survival analysis is based on protein expression in both in the overall cohort and dependent on the APC genetic background. This differences in survival are not as apparent in RNA data with the equivalent stratification.

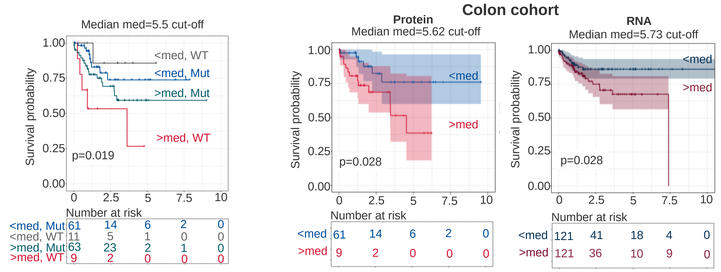

Figure 6. (A) Triaging patients based on DPEP1 log2 fold change (FC) median illustrates worse relapse-free survival for patients with higher protein expression FC in the tumor vs normal samples. (B) Analysis of the colon cancer subcohort shows worse survival with higher DPEP1 expression on both RNA and protein levels.

Therapeutic Potential

- Using multiple resources (CanSar, OpenTarget) we systematically defined top targets for their actionability using either antibody or small molecule based approaches.

Figure 7. (A) DPEP1 localizes to the nucleus, plasma membrane and extracellular space [1] and may hence be suitable for antibody therapy. (B) DPEP1 is a membrane enzyme, whose function may be drug targeted. (C) There exists an approved drug Cilastatin for DPEP1, which is used for bacterial infections and is being independently tested in CRC [1].

Conclusions

The IndivuType methodology successfully identifies potential drug targets for cancer. Here, we validate our method based on DPEP1, which was amongst the top targets in our selection. Indeed, DPEP1 has beneficial properties for targeting - convincingly, a drug targeting DPEP1 already exists. In collaboration with Evotec, we are currently testing other shortlisted targets experimentally without existing medication.

[1] Park *et al.*; Dehydropeptidase 1 promotes metastasis through regulation of E-cadherin expression in colon cancer., *Oncotarget*, 2016, Feb 23;7(8):9501-12, https://doi.org/10.18632/oncotarget.7033
[2] Binder *et al.;* COMPARTMENTS: unification and visualization of protein subcellular localization evidence, *Database*, Volume 2014, 2014, bau012, https://doi.org/10.1093/database/bau012

We have no potential conflict of interest to disclose.