

## **Project openings in the Molecular Cancer Genetics group at IGP, Rudbeck Laboratory**

*We are looking for students with an interest in genetic or functional studies of cancer, or in development of computer tools for large scale sequence data analysis.*

**Our research:** Mutations that cause normal cells to lose control over cell division and maintenance are important contributors to cancer development. Our main research topics are: (1) identification of cancer-causing mutations, (2) investigation of how specific mutations contribute to tumor development or metastasis, (3) strategies to utilize the specific genetic properties of cancer cells for targeted treatment, and (4) development of methods and procedures to aid or improve cancer diagnostics.

### **Current project proposals:**

1. Early detection and monitoring of cancer by analysis of circulating tumor DNA and other blood biomarkers has potential to reduce cancer mortality. However, several different types of analytes need to be determined in the one and same patient to achieve high sensitivity and specificity. Currently, this requires several different blood samples from the same patient processed in different and often intricate ways which limits clinical implementation. We are therefore developing a new pre-analytical technology to isolate several different fractions from the same blood sample which has potential to radically improve the implementation of liquid biopsies in the clinical workflow.

The work includes design, development and testing of an integrated system with hardware, software and consumables. The project has opportunities for several students with engineering competencies, such as molecular biotechnology, electrical engineering, programming of embedded systems, and mechanical engineering .

2. Hundreds of genes have been identified as cancer drivers following the introduction of next generation sequencing technologies in cancer genomics. Functional validation of candidate genes is essential to understand and utilize this information clinically. In a series of projects we have used rAAV gene targeting constructs to knock-out or knock-in function of candidate cancer genes in human cells to investigate their contribution to cancer associated phenotypes. We currently have an ongoing project for extensive characterization of different non-synonymous mutations of the oncogene *KRAS*. Isogenic cell lines are subjected to proteomic, transcriptomic and metabolomic profiling and the project will include follow-up of these analyses. Techniques that may be used include widely used methods such as cell culturing, PCR, RT-qPCR and Western blot.
3. We have recently reported non-synonymous mutations of *EPHB1* and their role in metastatic colon cancer (Mathot et al, *Cancer Research* 2017). As an extension to this project, we have screened for additional mutations of *EPHB1* and *EPHB2* in metastatic colon cancer using unique bioinformatics filtering. We currently have two possibilities for degree projects in this study:
  - I. Identified *EPHB1* and *EPHB2* mutations will be subject to validation *in vitro* by generation of isogenic cell models that are screened and characterized as described in

Mathot et al, 2017. The project will consist of extensive wet lab experimentation, involving e.g. cell culture, FACS sorting, RT-qPCR and Western blot.

- II. As a three dimensional structure, one should consider protein spatial coordinates when studying coding mutations in cancer. For that, we have been using X-Ray crystallography and nuclear magnetic resonance (NMR) protein structure files and a third party algorithm HotSpot3D (Niu et al, *Nature Genetics* 2016) to study related Ephrin protein family somatic mutations. Next in this project is the programming of a new algorithm based on the previous one but with improved logic and statistics. For this purpose we are looking for a student with basic knowledge in programming in Python. Experience with statistics and/or working with protein structures is also beneficial. The project includes the analysis of 3D protein structures by programming tasks with aim to output statistically significant 3D close proteins that are too far away in the linear structure to be considered as partners.
4. Targeted cancer therapy is based on finding conditions resulting in selective killing of cancer cells while sparing the normal tissues of the patient. We propose a concept based on exploitation of the natural genetic variation in the human population and the cancer specific phenomenon loss of heterozygosity (LOH) to identify tumors that are sensitized to certain drugs relative to the normal tissues. We have identified and ranked human candidate enzymes according to defined characteristics. For the candidate NAT2 we constructed and validated CRC cell model systems which in drug discovery efforts uncovered a compound with 3-fold increased cytotoxicity in cells lacking NAT2 *in vitro* and *in vivo*. Worldwide >50 000 colorectal cancer patients with NAT2 LOH could benefit from such treatment each year. In a similar effort we are now developing cell models for a second promising candidate target enzyme that will be studied during the course of this project. Techniques used in the project may include cell culturing, transfection, Western blot, PCR, functional cell analysis assays, and screening of chemical compound libraries.
5. We have developed software tools for rapid and accurate mutational analysis of deep sequencing data from solid tumors with significant content of normal cells. These tools have superior indel calling capabilities, a major challenge in mutational analysis, as compared to state of the art. For this application, novel statistical mathematics has been developed and patented (Swaminathan et al, *Pattern Recognition* 2016). The current challenge, pressed by ever tightening regulation on sequencing data management, is to assure the adequate storage and encryption of sensitive data. Therefore, we are looking for a highly motivated student to design or implement an existing encryption scheme, suitable for storage and processing of next-generation sequencing data. For the idea prototyping Python knowledge is a benefit, but other programming experience will be accepted. The technical side of the project includes the manipulation of data from several next-generation sequencing platforms and programming tasks based on the big data management.

For information and application contact Professor Tobias Sjöblom ([tobias.sjoblom@igp.uu.se](mailto:tobias.sjoblom@igp.uu.se)).