International PhD studies program: "MOLECULAR GENOMICS, TRANSCRIPTOMICS AND BIOINFORMATICS IN CANCER"

The future of discovery demands a group of talented scientists with unique training in both modern molecular biology and clinical medicine. **Molecular Genetics, Transcriptomics and Bioinformatics in Cancer** program will be constituted in the **Postgraduate School of Molecular Medicine**. The project will provide interdisciplinary postgraduate training focusing on application of recent high throughput technologies to study the molecular genetic and genomic processes central to molecular medicine. This project capitalizes on recent advances in our understanding of the molecular genetic mechanisms regulating cancer development, the bioinformatic analysis of genomic and postgenomic data, and proposes **an integrated, interdisciplinary approach.**

In the project transcriptomics/genomics data generated by analyses of follicular thyroid cancers, soft tissue sarcoma, melanoma, glioma, lung and breast cancers will be evaluated by dedicated bioinformatic methods in order to find characteristic gene signatures and/or novel predictive markers. In some systems transcriptomics and genome-wide data combined with sophisticated computational and system biology methods will be exploit to specify molecular pathways involved in carcinogenesis. Projects will be carried out in laboratories located: **in Warsaw** (Warsaw Medical University, M. Sklodowska-Curie Memorial, Cancer Center-Institute of Oncology, Interdisciplinary Centre for Mathematical and Computational Modelling of Warsaw University, Nencki Institute of Experimental Biology), **in Szczecin** (International Hereditary Cancer Centre, Pomeranian Medical University), **in Gliwice** (M. Sklodowska-Curie Memorial, Cancer Center-Institute of Oncology-Gliwice Branch). All teams have developed successful scientific cooperation in the proposed field of research (students will be working in collaborating laboratories for 6-12 months). Host institutes offer excellent working facilities with interactive groups and a competitive student stipend salary (3500 PLN/month).

The centerpiece of the project is a 3 course sequence (spring and summer of first year and the fall of second year) that provides in-depth training in the field of molecular genetic, genomic and bioinformatic research. Students will take a track-specific course "Genomics and Bioinformatics", "*Molecular biology of Cancer*" course emphasizing model systems where genetic regulation of disease are being understood and an integrated course "From laboratory bench to patient bed". All newly appointed students will be enrolled into The Postgraduate School of Molecular Medicine (SMM).

PhD project 1. Circulating tumor cells as prognostic and predictive factor in planning treatment of breast cancer - validation trial (supervisor prof. **Zbigniew Gaciong**, Medical University of Warsaw, Department of Internal Medicine, Hypertension and Vascular Diseases; collaborating scientist: Dr. **Piotr Religa**, Karolinska Institute, Stockholm, Sweden).

Recently, circulatory tumor cells (CTC) that create metastatic loci can be identified and counted in the peripheral blood. Presence and behavior of CTC provide information on the biology of cancer and factors such as structure of vessels, metastatic potential of tumor, their number correlate with patients survival at any time of breast cancer diseases. In patients with breast cancer CTC are identified based on expression of cytokeratin 19 along with other tumor markers by means of flow cytometry or RT-PCR. The main problem with clinical use of CTC as predictive and prognostic factor for planning of treatment is to lack of standards for quantitative estimation of cells and guidelines that can be used in clinical practice. Studies in the laboratory of Dr Religa delineated the role of progenitor cells in the process of neovascularization associated with tumor growth and formation of metastasis. Based on this experience and expertise we would like to apply similar approach to analyze the role of CTC in natural history of breast cancer in humans and experimental animals, asess prognostic and predictive value of CTC in progression of breast cancer and describe relationship between tumor biology and biology of CTC to determine if CTC are helpful factor in planning of breast cancer treatment. Implementation of new methods will allow future research goals: find potential new markers for counting of CTC; use the techniques for diagnosis other types of cancer such as: prostate, lung and colon cancer; find factors that could be helpful in treatment of cancer by modulation of structure of blood and lymphatic vessels. The project has two major parts: collection of patients' data and blood samples during neoadjuvant or palliative treatment, and detection of CTC in the collected material. RNA isolated from mononuclear cells will be used for cDNA and RTPCR to quantitate number of CTC. Following markers will be analyzed: cytokeratin 19, HER2/neu, ER, PR, mamoglobin. Isolation of DNA for estimation "tumor signature" is considered. **Two PhD student will be involved.**

PhD project 2. Molecular markers of thyroid cancer (supervisors **Barbara Jarząb/Andrzej Swierniak,** Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice Branch, Gliwice; collaborating scientist: **prof. Ralf Paschke,** Medical Clinic and III Polyclinic of Leipzig University, Germany)

The purpose of the study is to investigate mechanisms of neoplastic thyroid transformation with special emphasis on follicular thyroid carcinoma (FTC), the rare histotype much less known than the most frequent and intensively investigated type, papillary thyroid cancer (PTC). To broaden the understanding of FTC biology, we shall carry out gene expression profiling in malignant and benign follicular thyroid tumors by several approaches: (1) 3'-based oligonucleotide microarrays, targeted into 3' regions of each transcript (2) exon arrays (covering the whole measured transcript and allowing to detect splicing variants) (3) real-time quantitative PCR. The cooperation with Prof. Ralf Paschke and Dr Markus Eszlinger (Leipzig University) will result in evaluation of micro RNA profile (miRNA) in the same follicular tumors, malignant and benign, by a dedicated microRNA array.

Recently we have published new data on the role of miR-146a in the genetic predisposition to PTC (Jażdzewski et al. 2008, 2009). We plan to extend within the proposed project the analysis of different microRNAs in FTC and relate to its gene expression profile. The expression profile of follicular thyroid cancer will be compared to follicular adenomas, considered to be direct benign counterparts of FTC, as well as to a wider spectrum of other benign thyroid tumors, including polyclonal thyroid tumors , well characterized by their functional properties and histological features. Comparisons with PTC will be feasible as during

our previous studies we have gained an extensive amount of data on gene expression profile in this histotype.

Characterization of gene expression level will be extended by protein analysis (immunohistochemistry and Western blotting), to broader characterize and validate the observed differences and move forward the use of potential markers in clinical setting. Finally, information gained on molecular biology of FTC will be validated by cell lines-based assays, while the most characteristic expression molecular markers as well as miRNAs will be verified by quantitative real-time PCR, both in Germany and in Poland, not only in postoperative material but also in fine-needle biopsy specimens. For this goal, the PhD student will spend 6 months in Professor Paschke's/ Dr Eszlinger's laboratory.

The obtained genome-wide data will be evaluated by dedicated bioinformatical methods in order to find characteristic gene signatures as well as to specify molecular pathways involved in follicular thyroid carcinogenesis. The whole-gene coverage, offered by exon arrays will probably offer more reliable molecular markers: the extensive analysis of both approaches in diagnostic context will be carried out. The other direction of analysis will be the bioinformatical evaluation of reciprocal relations between array miRNA and array exon data. This step will also require a development of bioinformatical methods. Combined, our analysis shall provide a comprehensive evaluation of FTC biology at genomic level, both in the aspect of molecular mechanisms of this type of malignancy as well in the context of differential diagnosis. This will enable to propose new molecular FTC markers for their evaluation in further prospective clinical studies.

The project will be performed as **two parallel PhD projects, one devoted to biological aspects, using bioinformatical tools developed during the second PhD project by IT/biostastistics graduate.** As mentioned above, specific algorithm and methods have to be developed as both the commercial and freely available algorithms are not sufficient to finalize the necessary analyzes. This second PhD project is planned to be carried out within the shared PhD program (Faculty of Informatics and Systems Control, Silesian University of Technology). The bioinformatician will spend 6 months in the laboratory of Dr Knut Krohn, the bioinformatician cooperating with Professor Paschke at Leipzig University.

The continuation of this cooperation by the proposed project will enormously extend the knowledge about thyroid tumorigenesis by further unraveling the largely unknown molecular biology of follicular tumors and it will help to generate (mRNA and miRNA) markers for the differential diagnosis of FTC and FA (which will especially focus on the differential diagnosis of the non-specific FNAB entity "follicular proliferation").**Two PhD student will be involved.**

PhD project 3. Epigenetic modulation of genes involved in tumor and endometrial angiogenesis (supervisor Dr. **Pawel Wlodarski,** Medical University of Warsaw, collaborating scientist: Dr. **Claudine Kieda,** The Centre de Biophysique Moléculaire, Orleans, France)

Altered transcriptional activity of the gene in presence of required transcription factors is often related to either promoter polymorphism or its epigenetic silencing. Epigenetic modulation of gene expression includes histone modification (acetylation and deacetylation of lysine residues), gene silencing by promoter CpG island methylation or induction of microRNA.

The aim of this project is to determine epigenetic status of promoters of matrix metalloproteinases (MMP) and their inhibitors (TIMPs) in endometriosis and in neoplastic tumors of ovary.

Ovarian cancer is one of the leading causes of death. Besides well defined causative mutations (such as BRCA1/2) excessive angiogenesis is prominent. Endometriosis affects as much as 10% of women in the populations of developed countries. Endometriosis has a prevalence rate of 20-50% in infertile women and it is being considered one of the leading

causes of female infertility. Since angiogenesis accompanies development of the disease, it is reasoned that it may be critical for pathogenesis of endometriosis.
In the proposed project we intend to analyze methylation status of the promoters of TIMPs and MMPs genes in normal and diseased tissue. We will also determine levels of miRNA directed against TIMP. Results of these tests will be correlated with mRNA expression level of respective genes in the same samples. In addition to panel of diseased specimens, a number of analysis will be performed on the normal tissue. The latter one will be obtained during surgery on uterus or ovary conducted due to other causes. **One PhD student will be involved.**

PhD project 4. New prognostic and predictive markers for selected soft tissue sarcoma (supervisors **Piotr Rutkowski/Janusz A. Siedlecki,** M. Sklodowska-Curie Memorial, Cancer Center-Institute of Oncology, Warsaw)

Soft tissue sarcomas are the heterogeneous group of malignant mesenchymal tumors and they represent 1% of all neoplasms in adults. These tumors are characterized by poor clinical prognosis. Therefore, it is of a great importance to find new prognostic/predictive factors as well as new, possible therapeutic targets for these rare malicious tumors. We would like to focus our research on selected sarcoma types, such as synovial sarcoma, Ewing sarcoma, GIST and dermatofibrosarcoma. First aim is to find a new prognostic and predictive markers associated with long term survival of the selected sarcomas. Microarray technology will be applied to understand the biology of these tumors and to select genes which are differentially expressed in tumor versus control tissue. Second task will be to evaluate all the differentially expressed sarcoma specific genes as markers of the presence of sarcoma cells in routinely collected peripheral blood samples in specific sarcoma patients and try to correlate the result of RT-PCR assay with disease recurrence, survival and sarcoma prognostic factors. We hope that tissue specific markers can be used to detect the metastatic cells and at least be a markers of an earlier therapy. As some sarcomas are characterized by chromosome translocation, we would like to check if mRNA of such fused genes (e.g.: *SYT/SSX*) may served as a marker of synovial sarcoma cells circulating in peripheral blood. **One PhD student will be involved.**

PhD project 5. STAT-dependent transcriptional network and epigenetic modifications in melanomas and lung cancer cells (Supervisor **Bozena Kaminska,** Nencki Institute; collaborating scientist: Dr. **Salem Chouaib**, Institute Gustave Roussy, Villejuif, France)

Signal transducer and activator of transcription (STAT) are transcription factors phosphorylated in response to extracellular ligands that activate cytokine receptor signaling. Activated Stat3 participates in tumorigenesis by promoting uncontrolled growth and preventing apoptosis, and regulate immune evasion by inhibiting expression of proinflammatory mediators, while promoting expression of immune-suppressing factors. Dysregulation of Stat expression and/or signaling may lead to epigenetic changes resulting in global modification of gene expression, including repressive and stable epigenetic chromatin marks.

A major goal of this PhD project is identification of specific STAT dependent transcriptional network involved in altered tumor cell functions. Contribution of Stat1, 3, and 5 to genomic responses underlying tumor survival and progression melanoma and lung cancer models, will be characterized in collaboration with Dr. Chouaib's group. Using RNAi technology we will generate tumor cells stably depleted of specific Stat and global gene profiling will be analyzed to reveal putative Stat1, 3 and 5 co-regulated genes. Transcriptomic analysis of global pattern of gene expression and application of computational biology will allow to extract a valuable information on distinct gene sets and signaling pathways

dysregulated in cancer cells. To address possibility of epigenetic changes, global pattern of methylation, histon acetylation and methylation, will be determined in parental and given STAT-depleted cell lines using chromatin immunoprecipitation. Combining genome-wide detection *in vivo* protein-DNA interactions, maps of chromatin features and transcription factor localization, and application of computational methods will reveal gene regulatory mechanisms and networks underlying disease specific changes in gene expression. **One PhD students will be engaged to this project.**

PhD project 6. Genome-wide computational modeling of STAT transcriptional network in gliomas and brain inflammation (supervisor prof. **Jan Komorowski,** Interdisciplinary Centre for Mathematical and Computational Modelling, Warsaw; collaborating scientist: prof. **Claes Wadelius**, Uppsala University, Sweden)

Gliomas are the most common primary brain tumors characterized by infiltration of tumor cells into the brain tissue contributes to the failure of current treatments and poor survival of patients. Recent studies demonstrated the occurrence of tumor associated inflammation contributing to glioma growth and invasion. An excessive or prolonged inflammation is deleterious and this process significantly contribute to the progression of cancer and several major neurodegenerative diseases. A crucial role of the STAT family in coordinating genomic responses in inflammation and cancer. Although these TFs bind to highly similar or even overlapping sequences, the net effect of these interactions will differ depending on which of STAT members are bound under a given condition to a particular gene. Using the **Ultra High Throughput (UHT) sequencing**, we will provide a global view of interactions between STAT1, STAT3 and STAT5 and the genome in glioma cells and cell culture model of inflammation. A major scientific goal for this project is to contribute with computational methods and tools that enable genome-wide analysis of interactions between transcriptional regulators and DNA using data generated by UHT sequencing. In collaboration with Professor Wadelius, Uppsala University, novel tools for the analysis of UHT data will be developed and applied. Experience in bioinformatics or computer science required. **One PhD student will be involved.**

PhD project 7. Development of novel algorithms for bioinformatic analysis of cancer-related signaling pathways (supervisor **Lucjan Wyrwicz,** M. Sklodowska-Curie Memorial, Cancer Center-Institute of Oncology, Warsaw), collaborating scientist: Dr. **Alexander Kel** (BIOBASE Biological Databases, Wolfenbuettel, Germany)

The inventions in fields of microarrays and proteomics are a great source of data on behavior of molecular systems. With great expectations coming from high throughput molecular technologies, the data analysis still lacks the reference methods aimed on modeling of signaling pathways. Since pathogenesis of cancer is either directly or indirectly related to disturbances in cell signaling – the most important task for cancer computational biology is the introduction of novel methods to understand cancer-related alterations in cell. The functional analysis of high throughput experiments relies on the description of genes' function. With a relatively high degree of confidence in knowledge on molecular function of genes, there is still a need for methods aimed on classification of biological role of gene products in cell.

In this project we plan to develop a method for reconstruction of **new significant functional associations of genes by assembly of various sources of data**. The data will be derived from various sources – including sequence, structure, expression, interactions and regulation of given gene or its protein product. The background of this task follow the postulate that functionally related genes can be characterized with some common features, that find a reflection in the knowledge deposited in public repositories, like literature databases, results of

biological experiments and bioinformatical predictions. The predictions will be assessed in a proper statistical tests, validated by reconstruction of known signaling pathways and, if possible, it will be confirmed by collaborating experimental biologists. The obtained “gene sets” will be further on tested in analysis of microarrays and proteomics experiments in typical “gene set enrichment analyses. **One PhD student will be involved.**

PhD project 8. Influence of modifier genes on the risk of breast and/or ovarian cancer In *BRCA1* mutation carriers. (supervisors **Jan Lubiński/Anna Jakubowska**, International Hereditary Cancer Centre, Pomeranian Medical University, Szczecin)

The major inherited risks of development of breast and ovarian cancers are germline mutations in either the *BRCA1* or *BRCA2* genes. The substantial variability in breast cancer penetrance in *BRCA1* mutation carriers is probably explained by modification by other genetic or environmental factors. The knowledge of risk-modifying factors in addition to *BRCA1* mutation status would aid in the effective application of risk prediction or cancer prevention strategies in women carrying *BRCA1* mutations. In Poland the three common founder mutations in *BRCA1* gene account for the majority (~ 90%) of *BRCA1* mutations in breast-ovarian cancer families making it is ideal for association studies of risk of cancer development and progression modifying genes not influenced by *BRCA1* allelic or ethnic variations. The aim of this study is to search for breast and ovarian cancer susceptibility genes that modify the breast and ovarian cancer risks in women carrying one of the three common *BRCA1* founder mutations. We will analyze 1000 women affected by breast or ovarian cancer as well as unaffected carriers, the results showing an association of particular genes/SNPs with *BRCA1* related breast/ovarian cancer risk will be verified on a large group of carriers (~15,000) from **CIMBA** (Consortium of Investigators of Modifiers of BRCA1 and BRCA2; the multicenter collaborative study). **Two PhD students will be involved.**