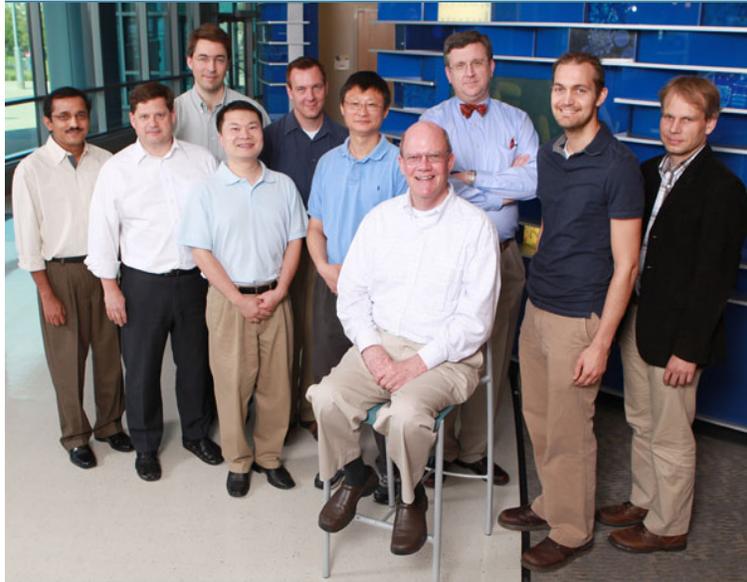## Division Photo

Front: John Hutton (sitting), Keith Marsolo, Michal Kouril.
2nd row: Anil Jegga, John Pestian, Jason Lu, Jun Ma, Andy Spooner
Back: Michael Wagner, Eric Hall

Division Data Summary

Research and Training Details

Number of Faculty	10
Number of Joint Appointment Faculty	5
Number of Research Fellows	4
Number of Research Students	12
Number of Support Personnel	60
Direct Annual Grant Support	$1,720,833
Direct Annual Industry Support	$78,261
Peer Reviewed Publications	27

Significant Publications

Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 2009 Jul 1;37(Web Server issue):W305-11

ToppGene Suite (http://toppgene.cchmc.org) is an in-house developed, widely-used, user-friendly, one-stop portal for performing (i) gene list functional enrichment analysis, (ii) candidate gene prioritization using either functional annotations or network analysis and (iii) identification and prioritization of novel disease candidate genes in the protein interactome space. Existing disease candidate gene prioritization methodologies mine biological and functional information about candidate genes, and ToppGene Suite can complement these existing approaches by applying novel methods that mine mouse phenotype data and protein interactions. We strongly believe that it will be a valuable adjunct to wet lab experiments in human genetics and disease research.

ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems Kaimal V, Bardes EE, Tabar SC, Jegga AG, Aronow BJ Nucleic Acids Res. 2010 Jul 1;38 Suppl:W96-102. Epub 2010 May 19 PMID: 20484371

The manuscript "ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems" is a breakthrough for what many in the field of bioionformatics try to do for biomedical research. It represents a new kind of genome function analysis tool that gives biomedical researchers an unprecedented capability to explore and discover key biological mechanisms responsible for many aspects of health and disease. The ToppCluster server is an online and freely accessible analysis system that allows a researcher to enter lists of genes from their own studies, or from those of others, and carry out knowledge analyses at a deeper level than has been previously possible in a single data mining environment. Currently it uses 17-dimensions of gene-associated knowledge, including thousands of manuscripts and findings from biomedical literature, protein-protein interactions, and the carefully tabulated positive and negative effects of gene function and gene dysfunction and provides researchers the ability to build state of the art graph network models of biological systems. The resulting graph networks allow researchers to compare and contrast massive amounts of gene-associated knowledge and data, and identify relationships that provide a higher level of detail than has previously been possible.

Division Highlights

Using Computational Biology to Study Disease

Drs. Bruce Aronow, Anil Jegga, and Jason Lu focus on the study of human disease mechanisms using computational approaches. Initiated under the aegis of the Computational Medicine Center and supported by numerous center grants including the UC/CHMC CTSA, Jegga and Aronow have developed a unique web portal for the analysis of disease biology data and knowledge. These next-generation research tools, called ToppGene.cchmc.org, ToppCluster.cchmc.org, and GATACA.cchmc.org, enable biomedical researchers to explore and predict pathways responsible for disease causation, treatment, and prevention and have contributed to numerous recent grants, publications, and new research studies. As leading contributors to the field of Disease Systems Biology, this group effort of BMI has provided both Cincinnati and international collaborations with exciting new directions and insights. The recently published tool ToppCluster provides a basis for researchers to combine diverse knowledge and clinical and laboratory data sets in such a way as to gain a deeper ability to model pathways and mechanisms responsible for how systems function or undergo disease pathology. Critical for the accomplishment of this was the recently developed comparative mouse models vs. human disease phenomics database ("PhenoHM") that combines and extrapolates knowledge about the roles of individual gene functions in the determination of phenotype across multiple organisms. The Phenome.cchmc.org server provides a resource that enables rapid identification of genes that trigger similar outcomes in human and mouse and facilitates identification of potentially novel disease causal genes. Lu focuses on using network/systems approaches to study the fundamental problems underlying human diseases. In the area of infectious diseases, his group has developed a novel computational algorithm that allows rapid and accurate predictions of essential genes in unknown bacterial pathogens. These genes are essential to the viability of the pathogen and are the most obvious targets for new drugs designed to kill the organism and to treat infection. In the area of cardiovascular diseases (CVD), his group has set out to discover the molecular basis underlying the High Density Lipoprotein (HDL) heterogeneity which could be a basis for therapeutic strategies.

Optimizing IT support for research

In his role as Faculty Liasison, Dr. Michael Wagner reaches out to CHRF faculty and works with BMI and IS staff to optimize IT support of their research programs. Michael is the informatics lead on both the P30 Cincinnati Rheumatic Diseases Core Center (Susan Thompson, PI) and the Pediatric Functional Neuroimaging Research Network (Scott Holland, PI), and oversees the development of software infrastructure needs for these two large centers. Along with the streamlined web portals that facilitate primary research data management and processing for the flow, microarray and genotyping cores, BMI staff are now implementing a web-facing reference database and cluster-based processing pipeline for pediatric fMRI data. Furthermore, the web-based core facility scheduling software which was deployed in the last year has now been augmented by a time tracking and billing module which greatly facilitates the accuracy and transparency of the billing processes for the flow and confocal microscopy cores.

Clinical Linguistics

The Clinical Linguistic group, led by Dr. John Pestian, focuses on a novel research and development that integrates clinical text with the various biological correlates to develop methods for disease specific diagnosis and treatments. This is being done in collaboration with faculty from the Divisions of Psychiatry, Emergency Medicine, Neurology and Hematology/Oncology. Research and development efforts include the development of the worlds largest corpus of suicide notes, pilot testing of methods for understanding the subjective language of suicidal and end-of-life patients, quality measures for epilepsy treatment, and drug selection for neuro-psychiatric disorders.

Other programs

Other programs include Dr. Jun Ma's research group that focuses on computational modeling of developmental processes in Drosophila and research of affiliated faculty focusing on protein modeling (Jarek Meller), statistical modeling of biological processes (Mario Medvedovic), and gene regulatory networks in the lung. Dr. Andy Spooner, Associate Professor in the Division, is Chief Medical Officer of CCHMC and has research interests in clinical information systems and decision support.

Graduate Education

The graduate program in bioinformatics is based in the UC College of Engineering, where the CCHMC biomedical informatics' faculty have joint appointments. Dr. Jarek Meller leads the program in close collaboration with other faculty

from Biomedical Informatics (Drs. Aronow, Jegga, Lu, and Medvedovic). Currently fifteen PhD candidates are enrolled and 2 new students are to be admitted in the Fall, 2010. Courses offered include Introduction to Functional Genomics, Introduction to Bioinformatics, and Advanced Computational Biology. Interdisciplinary workshops and multimedia educational materials are offered for medical students and other professionals, building on our cross-institutional efforts in the form of seminar series and journal club, summer workshops and other training initiatives. Thanks to the strength of the program, several Choose Ohio First scholarships have been secured this year for our current and prospective students.

Division Collaboration

Collaboration with Allergy & Immunology; Gastroenterology, Hepatology, and Nutrition

Collaborating Faculty: Pabo Abonia; James Franciosi; Marc Rothenberg

Dr. Keith Marsolo and the i2b2 team are leading the development of a multi-center registry focused on Eosinophilic Esophagitis. This registry takes advantage of functionality developed for Liver Transplant and also allows extracts from other sites to be uploaded into the registry.

Collaboration with Behavioral Medicine

Collaborating Faculty: Jennie Noll

Dr. Michal Kouril (BMI) worked with Dr. Jennie Noll on setting up a platform for assessment of online behavior of high risk female adolescents.

Collaboration with Developmental Biology

Collaborating Faculty: James Lessard; Steven Potter

Dr. Bruce Aronow's group collaborates with Potter and Lessard along with an international consortium on the use of genomics analyses to gain insight into the normal or abnormal development of the kidney and lower urinary tract.

Collaboration with Experimental Hematology and Cancer Biology

Collaborating Faculty: Nancy Ratner

Dr. Bruce Aronow and his team's collaborative work with Dr. Ratner involves comparative genomics analyses of human and mouse developing and oncogenically transformed nervous system cells and tissues and using this information to predict novel or optimal therapeutic agents and gene-based targets.

Collaboration with Gastroenterology, Hepatology, and Nutrition

Collaborating Faculty: Jorge Bezerra

Dr. Anil Jegga collaborates with Bezerra Lab to understand the molecular basis of biliary atresia, a rare condition in newborn infants which if unrecognized could lead to liver failure. He provides bioinformatic support that includes data analysis, data-mining and hypothesis generation using systems biology-based approaches.

Collaboration with Imaging Research Center

Collaborating Faculty: Scott Holland; Vince Schmithorst; Jennifer Vannest

Dr. Michael Wagner's collaboration with the Imaging Research Center is focused on building a pediatric fMRI reference database of normal brain development and is funded by a large NIH contract.

Collaboration with Immunobiology

Collaborating Faculty: Lee Grimes

Dr. Anil Jegga collaborates with Grimes Lab in their continuing pursuits to characterize cancer proteins and understand the underlying regulatory mechanisms of oncogenic transformation of hematopoietic progenitor cells.

Collaboration with Neurology

Collaborating Faculty: Tracy Glauser

Dr. John Pestian and his team are collaborating with Dr. Glauser for the on going development of CHRISTINE, a clinical decision support system for identifying optimal drug therapy for patients with epilepsy, and ADHD.

Collaboration with Physical Medicine and Rehabilitation

Collaborating Faculty: Shari Wade

Dr. Keith Marsolo and his team are implementing web-based interventions for patients and families as part of the Rehabilitation Research and Training Center grant.

Collaboration with Psychiatry; Emergency Medicine

Collaborating Faculty: Robert Kowatch; Jacqueline Grupp-Phelan

Dr. John Pestian and his team are collaborating with the Psychiatry and Emergency Medicine Divisions to develop methods to identify the likelihood of repeated suicide attempts.

Collaboration with Rheumatology

Collaborating Faculty: Sue Thompson

Dr. Michael Wagner is collaborating with Sue Thompson on a large project focusing on integrative genetics in JIA (funding comes from a P30 Center grant and other Rheumatology grants)

Faculty Members

John Hutton, MD, Professor ; *Director, Division Chief*

Bruce Aronow, PhD, Professor ; *Co-Director, Computational Medicine Center*
   Research Interests: Gene Expression Analysis, Gene Regulation, Clinical Genomics, Functional Genomics of Development and Disease

Anil Jegga, MS, DVM, Research Assistant Professor
  Research Interests: Gene Regulatory Networks, Biomedical Ontologies, Integrative Genomics

Michal Kouril, PhD, Research Instructor ; *Director, Research IT*
  Research Interests: Computational Support, High-performance computing, Parallel Programming, High-end Data Storage

Long Jason Lu, PhD, Assistant Professor
  Research Interests: Bioinformatics, Machine Learning, Integrative Genomics, Biological Networks, Computational Modeling, Software Development

Jun Ma, PhD, Associate Professor
  Research Interests: Development, Transcription, Morphogen Gradient, Embryo, Robustness, Quantitative Studies

Keith Marsolo, PhD, Research Instructor ; *Director, Software Development and Data Warehouse*
  Research Interests: i2b2, Data Integration, Data Warehousing and Data Management

John Pestian, PhD, MBA, Associate Professor ; *Director, Computational Medicine Center*
  Research Interests: Natural Language Processing, Clinical Decision Support, Suicide Research, Pathology Research, Psychiatric Research

S. Andrew Spooner, MD, FAAP, Associate Professor Clinical ; *Chief Medical Information Officer*
  Research Interests: Decision Support, Pharmacy Information Systems

Michael Wagner, PhD, Research Associate Professor ; *Faculty Liaison*
  Research Interests: Machine Learning, Proteomics, Genome-wide Association, Parallel Computing, Computational Infrastructure, Bioinformatics

Joint Appointment Faculty Members

Karen Davis, PhD, Associate Professor
  UC ECECS
  Electrical & Computer Engineering and Computer Science

Eric Hall, PhD, Research Instructor
  Neonatology & Pulmonary Biology
  Clinical Informatics, Knowledge Discovery Tools, Data Mining and Warehousing

Mario Medvedovic, PhD, Associate Professor
  UC Environmental Health
  Biostatistics

Jarek Meller, PhD, Associate Professor
  UC Environmental Health
  Protein Modeling

Yan Xu, PhD, Associate Professor
  Pulmonary Medicine

Trainees

- **Jacek Biesiada, PhD,** University of Silesia, Poland
- **Junbo Liu, PhD,** Fudan University, Shanghai, China
- **Julia Taylor, PhD,** University of Cincinnati, Cincinnati, OH

Significant Accomplishments

Our mission

Our Division provides the core knowledge, skills and equipment necessary to acquire, store and analyze biomedical data using computational systems. We work with medical researchers to incorporate this technology into their work.

Research IT services

The Research Data Center, led by Michael Kouril, PhD, serves as liaison between Biomedical Informatics and technical groups in Information Services at Cincinnati Children's and the University of Cincinnati. Kouril's team works to improve the speed of communication between the institutions while maintaining strong data security.

Data management and software development

The data warehousing and software development group, headed by Keith Marsolo, PhD, is expanding the framework of our research patient data warehouse to serve as a platform for research registries. When combined with distributed query technologies, this platform can support multi-center registries for managing chronic disease, monitoring product and drug safety, tracking the natural history of disease, and improving quality. It is being piloted in two NIH-funded projects. The

first, based at Cincinnati Children's, will be a multi-center registry focused on eosinophilic esophagitis. The second project, with Children's Hospital Boston, will create a registry of more than 60 sites targeted to pediatric arthritis and other rheumatologic diseases.

Eric Hall, PhD, who works jointly with our division and Neonatology, is working with Marsolo to integrate geographic data on preterm births and the distribution of risk factors into the platform, allowing targeted interventions to the most at-risk populations. The team also is establishing a secure, confidential regional perinatal data repository to better recognize and manage neonatal and childhood diseases.

Additionally, they are helping investigators in intestinal rehabilitation to improve care and research by developing a patient registry integrated with the electronic health record.

Division Publications

1. :

Grants, Contracts, and Industry Agreements

Grant and Contract Awards		Annual Direct / Project Period Direct
Aronow, B		
Molecular Signatures of Cancer Metastasis		
Department of Defense		
W81XWH-10-1-0325	05/01/10 - 04/30/13	$99,997 / $299,997
Risk Stratification & Identification of Immunogenetic and Microbial Markers of Complicated Disease Course in Pediatric Crohn's Disease		
Emory University (Crohn's and Colitis Foundation of America)		
	07/09/09 - 06/30/13	$12,116 / $51,552
Exploiting Advances in Biotechnology for Force Protection		
United States Air Force (University of Cincinnati)		
FA865008C6832	07/01/09 - 06/30/14	$264,027 / $1,320,135
Cincinnati Center for Clinical & Translational Sciences and Training - Translational and Molecular Disease Modeling		
University of Cincinnati (National Institutes of Health)		
UL1 RR 026314	04/03/09 - 03/31/14	$31,502 / $55,740
Hutton, J		
Enterprise Research Data Storage for Data-Intensive Computation		
National Center for Research Resources		
S10 RR 027230	02/20/10 - 02/19/11	$500,000 / $500,000
Jegga, A		
Functional Polymorphisms in p53 Response Elements		
Ohio Cancer Research Associates		
	07/01/08 - 12/31/10	$22,727 / $45,454
Kouril/Marsolo/Wagner		
Cincinnati Center for Clinical and Translational Sciences and Training - Biomedical Informatics		
University of Cincinnati (National Institutes of Health)		
UL1 RR 026314	04/03/09 - 03/31/14	$84,254 / $156,429
Lu, L		
An Integrative Approach for Predicting Essential Genes		
University of Cincinnati		
	09/15/09 - 09/14/10	$15,557 / $15,557
Ma, J		
Probing the Robustness of a Developmental System		
National Science Foundation		
IOS-0843424	05/15/09 - 04/30/13	$197,419 / $855,599
Evolution of Canalizing Mechanisms in Gene Expression		

The University of Chicago (National Institutes of Health)
R01 GM 78381   07/16/09 - 06/30/11 $74,528 / $74,528

Marsolo, K

CARRA: Accelerating Toward an Evidence Based Culture in Pediatric Rheumatology
Duke University (National Institutes of Health)
RC2 AR 058934   09/30/09 - 08/31/11 $282,697 / $473,120

Spooner, A

Step Stools: Developing Web Services for Safe Pediatric Dosing
Vanderbilt University Medical Center (National Institutes of Health)
R18 HS 017216   09/01/08 - 08/31/10 $24,981 / $37,049

Wagner, M

Bayesian Mixture Modeling of Functional Genomics Data
University of Cincinnati (National Institutes of Health)
R01 HG 003749   07/01/08 - 06/30/10 $14,521 / $53,534

Optimization of Structures and Networks of Proteins
The University of Texas at Austin (National Institutes of Health)
R01 GM 067823   06/01/08 - 05/31/11 $39,410 / $57,286

Cincinnati Rheumatic Diseases Core Center - Core 4
National Institutes of Health
P30 AR 047363   07/01/09 - 06/30/10 $57,097 / $57,097

	Current Year Direct	**$ 1,720,833**

Industry Contracts

Pestian, J
EISAI Europe, Ltd. $ 78,261

	Current Year Direct Receipts	**$ 78,261**

	Total	**$1,799,094**