

Background

FOR OFFICE USE ONLY: Version # _____ APP # _____

Applicant Information

(Carefully read the instructions before completing this form)

1. Applicant Information

a. Applicant Name	Anne-Claude Gingras	Title	<input checked="" type="radio"/> Dr	<input type="radio"/> Mr	<input type="radio"/> Ms	<input type="radio"/> Prof
b. Institution	Mount Sinai Hospital					
c. Department						
d. Mailing Address	Samuel Lunenfeld Research Institute					
e. Address 2	Mount Sinai Hospital					
f. Address 3	600 University Avenue, Suite 992A					
g. Address 4						
h. Address 5					Country	CA
i. City	Toronto	Province	ON	Postal Code	M5G 1X5	
j. Phone	(416) 586 4800	Ext	5027	Fax	(416) 586 8857	
k. e-Mail Address	gingras@lunenfeld.ca					
l. Designation	Associate Professor					

2. Project Information

a. Project Title	Rapid mapping of cancer interactomes				
b. Is Financial Institution the same as the Research Institution? (Please select Yes or No)			<input checked="" type="radio"/> Yes	<input type="radio"/> No	
c. If No, provide Financial Institution name					
d. Project Start Date	Feb-01-2013	End Date	Jan-31-2015		
e. Amount of Funds Requested	\$199,994.00	Project Cost	\$199,994.00		
f. Grant category (Please select one from list)					
	<input checked="" type="radio"/> Research Grant				
	<input type="radio"/> New Investigator				
g. Type of application. Note: maximum 1 application allowed per competition (Please select one from list)					
	<input checked="" type="radio"/> Initial Application				
	<input type="radio"/> Re-application				
h. The applicants confirm (by choosing "Yes") that as of October 1, 2012 there is no substantive (more than 50%) overlap with any other pending application		Yes			
i. Indicate the number of years of support requested (term can extend over 3 years)	2				
j. Is this application being submitted in French? (Please select Yes or No)		<input type="radio"/> Yes	<input checked="" type="radio"/> No		

FOR OFFICE USE ONLY: Version # _____

APP # _____

3. Participants

a. Financial Officer

Title	Mr		
Name	Paul Kranjac		
Department	Director of Finance		
Institution	Mount Sinai Hospital		
Address 1	600 University Avenue		
Address 2			
Address 3			
Country	CA		
City	Toronto	Province ON	Postal Code M5G 1X5
Telephone	(416) 586-5398	Fax	(416) 586-8588
E-Mail Address	kranjac@lunenfeld.ca		
Title	Financial Manager		

b. Co-Applicant

Title	Prof		
Name	Philip Kim		
Department	160 College St., Rm 606		
Institution	CCBR		
Address 1	University of Toronto		
Address 2			
Address 3			
Country	CA		
City	Toronto	Province ON	Postal Code M5S 3E1
Telephone	(416) 946-3419	Fax	
E-Mail Address	pm.kim@utoronto.ca		
Title	Professor		
Curriculum Vitae (CV)	Kim_philip-CV.pdf		

Applicant info

FOR OFFICE USE ONLY: Version # _____ APP # _____

4. Principal Investigator CV

Attachments:

[Gingras_anneclaude-CV](#)

5. Letters of collaboration

Attachments:

[gingras_anneclaude-collaborations](#)

6. Stage of career development (optional)

Please indicate below if this is:

Your first application for a research grant to the Canadian Cancer Society Research Institute (Please select Yes or No) Yes No

Your first application for a research grant specifically in the area of cancer research (Please select Yes or No) Yes No

Your first application for a research grant as an independent investigator (Please select Yes or No) Yes No

7. Biographical information (optional)

Gender

Male

Female

Date of birth - month 4 Day 10 Year 1972

Birth city Quebec City Birth province QC Birth country Canada

City raised Ste-Petronille Country raised Canada

Language(s) spoken Français, English

Citizenship CANADIAN

Certificates

FOR OFFICE USE ONLY:	Version # _____	APP # _____
----------------------	-----------------	-------------

8. Certificates required

8.a1 Biohazard/Biosafety

- a. Does your project require a biohazard certificate? (Please select Yes or No) Yes No
- b. If required, state the status of your certificate
 Included with application To follow
- c. Are biohazard certificates required from other institutions (in case of Co-Principal Investigators and Co-Applicants)? (Please select Yes or No) Yes No
- d. If yes, give details.

Institution	Status
-------------	--------

8.a2 Biohazard/Biosafety certificate

8.a3 Biohazard/Biosafety certificates (other institutions)

8.b1 Animal care

- a. Does your project require animal care certificates? (Please select Yes or No) Yes No
- b. If required, indicate the status:
 Included with application To follow
- c. Are animal care certificates required from other institutions (in case of Co-Principal Investigators and Co-Applicants)? (Please select Yes or No) Yes No
- d. If yes, give details.

Institution	Status
-------------	--------

8.b2 Animal care certificate

8.b3 Animal care certificates (other institutions)

8.c1 Ethics

- a. Does your project require ethics certificates? (Please select Yes or No) Yes No
- b. If required, indicate the status:
 Included with application To follow
- c. Are ethics certificates required from other institutions (in case of Co-Principal Investigators and Co-Applicants)? (Please select Yes or No) Yes No
- d. If yes, give details.

Institution	Status
-------------	--------

8.c2 Ethics certificate

8.c3 Ethics certificate (other institutions)

9. Human embryonic stem cells involvement

Does the proposal involve the use or creation of human embryonic stem cells? (If yes, contact the CCSRI) (Please select Yes or No)

Yes

No

If yes, is the research reviewed under the auspices of the local ethics review board? (Do not answer this if the answer above is No)

Public summary

FOR OFFICE USE ONLY:

Version # _____

APP # _____

10. Public summary - project summary

Proteins inside each cell bind to each other to perform their activities. When mutations in a protein arise, many of these binding events are altered, and these can result in cancer. Recently, the scientific community has created excellent methods to identify the mutations, but this has thus far not been matched by analyses at the molecular level, including such binding events. Here, we are proposing to use innovative tools to systematically identify those binding events that are modulated after a mutation occurs. These tools should help bridge the gap between the descriptive information (what are the mutations?) provided by the sequencing of the genetic code, and functional information (what does this mutation do?).

11. Public summary - previous research

The Gingras lab has been developing new approaches to accurately quantify each of the proteins present in a biological sample using a machine called a mass spectrometer that can essentially identify proteins based on "fingerprints". They are also recognized internationally for their studies of binding events between proteins. They have shown on a small test case that a new approach they developed can help in identifying how specific mutations can affect cancer development by changing how proteins bind to each other. Through collaborations, they now have access to large a collection of mutations identified in cancer cells. The Kim lab's expertise is in defining how proteins interact with each other in a very precise manner. They develop computational tools which enable them to predict how a mutation could influence the binding of proteins to one another.

12. Public summary - project description

The Gingras and Kim labs will join efforts to provide a method to systematically identify changes in protein-protein interactions upon expression of proteins containing cancer associated mutations. They will first computationally predict which mutations are most likely to affect protein-protein interactions. Then, using new technology being developed in the Gingras laboratory, they will systematically test these; for this, they will use ~150 different mutant proteins. They will also predict which of the interactions can be disrupted by drugs, and test these predictions experimentally. What they are expecting to generate is both new biology on the mutations that they are studying that may explain how cancer arises, but also a methodology to systematically assess many more mutations in the future. They will distribute their results and methods to the scientific community, accelerating the discovery of the causes of cancer and of treatment options.

13. Public summary - impact and relevance statement

Understanding how a mutation causes cancer is critical to defining treatment options. For example, some drugs act by preventing the association of proteins with one another. This is the case of a drug that the Gingras lab used in a pilot study: in this case, they demonstrated that the mutant was resistant to the effects of this drug, which is important to know when treating patients that present with this particular mutation. The ultimate goal of this project is to accelerate the discovery of drug targets for cancers by providing more information about the mechanistic aspects of cancer development following protein mutation. While this is a long-term goal, immediate consequences on the life of Canadians living with cancers may result from one of our aims, which is to test known drugs for the changes they induce in protein binding events for non-mutated and mutated forms of proteins.

Abstract

FOR OFFICE USE ONLY:

Version # _____

APP # _____

14. Scientific abstract

With the advent of next generation DNA sequencing technologies, the pace of discovery of cancer-associated sequence variants has greatly accelerated, leading to the realization that tumours (especially solid tumours) harbour hundreds of mutations. Ongoing statistical analysis across multiple laboratories worldwide is progressing rapidly, helping to identify which of these mutations are likely drivers of the cancer phenotype. However, in spite of rapid progress in mapping of cancer related signaling and interaction networks, there has been an increasing disconnect between the identification of a cancer variant, and the mechanistic understanding of transformation induced by this mutation. While the downstream effects of many of the mutations may be straightforward to explain (e.g., abrogation of protein expression or function for tumour suppressors), what happens at the molecular level as a consequence of a point mutation is generally not well understood. Such molecular understanding is crucial for developing therapeutic interventions. Frequent consequences of cancer mutations (especially point mutations in the coding sequence) are specific alterations of protein-protein interactions. As we are developing both appropriate structural data and technology to assess and modulate such interactions, an analysis of specific protein-protein interactions altered by cancer-associated mutations is timely and could offer novel therapeutic avenues.

To ideally exploit this novel venue, we should be able to quantitatively measure both loss and gain of interactions, be able to rapidly measure different sequence variants, and have a robust, sensitive and accurate approach. In order to develop such a methodology, we have coupled affinity purification (AP) to a novel data-independent mass spectrometric acquisition (SWATH). We have then tested this strategy for the quantification of the changes imparted by melanoma-associated mutations in the kinase CDK4. In addition to confirming known interaction loss, we were able to also identify interaction gains in the same experiment, namely in the recruitment of HSP90 network components to the mutated proteins. We next used AP-SWATH to monitor the consequence of pharmacological inhibition of HSP90 on the recruitment of chaperones to the wild type and mutated CDK4 proteins and rapidly demonstrated that the mutant proteins are resistant to the inhibitor. This finding has obvious implications for patient care.

Based on these results, we are proposing to develop an experimental and computational pipeline to assess interactions changes. We will make this pipeline scalable to optimally make use of ongoing genomic sequencing efforts. We have access through our collaborations to a large collection of disease alleles for human proteins that will be prioritized for analysis by AP-SWATH based on the structure-based predictions of mutations most likely to disrupt protein-protein interactions (this is the area of expertise of co-applicant Kim). In Aim 1, selected mutants (and their wild type counterparts ~100) will be analyzed by AP-SWATH to identify deregulated interactions. This will be followed in Aim 2 by a deeper analysis of the changes imparted in interactions for phosphatases (an intense area of research in the Gingras lab). In Aim 3, we will explore the consequences of treatment with pharmacological inhibitors on the interactions for wild type and mutant proteins.

Throughout the duration of the grant, we will optimize the platform to enable higher throughput analysis and provide cancer researchers with annotation of interactome changes imparted by mutations, and when possible, by the treatment with pharmaceuticals. In addition, we will be addressing yet unresolved questions: for how many (and which) of the point mutations are interaction profiles altered and how should the community embark on a systematic characterization of these changes? Our project will greatly enhance our

understanding of the effects of cancer mutations both on a cellular and molecular level, and provide the research community with leads for biological follow-up, thereby accelerating the development of therapeutic options.

15. Keywords/Technical terms

Keyword/Technical terms
mutation
mass spectrometry
affinity purification
edgetic mutation
SWATH (Data independent acquisition technology)
protein-protein interaction
data analysis
structural prediction
cancer drug

16. Innovation statement

The pace of discovery of cancer-associated mutations has been accelerating greatly in the past years, aided by major advances in DNA sequencing technologies, but the functional consequences of the great majority of these mutations are poorly understood. Several years ago, our collaborators (Vidal/Hill) began generating a collection of constructs encoding wild type and mutated alleles for proteins implicated in many diseases, including cancer, and shown using yeast-two hybrid technologies that several of these modulate specific protein-protein interactions.

The innovative aspects of our project rest on the development of new mass spectrometric approaches to accurately quantify altered protein-protein interactions in a scalable platform (Gingras), and the use of existing three-dimensional data and computational modeling to focus on mutations likely to affect interactions (Kim).

We have been working for many years on the development of affinity purification coupled with mass spectrometry to identify protein-protein interactions. While our previous approaches were appropriate for identifying static interactions, they largely failed at identifying differentially regulated interactions. A new concept in quantitative mass spectrometry, termed SWATH, enables the label-free quantification of all identifiable species in the mass spectrometer. Though it is very new (first publication in March 2012), the Lunenfeld has been an early adopter of the method, which we have optimized extensively. Because of its potential for monitoring interactome changes in a rapid, unbiased and accurate manner, we have coupled affinity purification (AP) to SWATH and validated the approach on cancer-associated mutations. Here, we will develop this approach into a scalable platform which will enable cancer biologists to rapidly identify interactions and pathways affected by cancer-associated mutations. This will be done in collaboration with AB-SCIEX and a number of computational experts, enabling us to create an innovative pipeline for functional analysis of cancer mutations.

Proposal

FOR OFFICE USE ONLY:

Version # _____

APP # _____

17. Abstract changes

- a. The CCSRI has done an initial review of your abstract. Please indicate if significant revisions have occurred since its submission. (Please select Yes or No)

Yes

No

18. Response to previous critique (re-applications only)

18-a. Response to previous critique (re-applications only)

18-b. Scientific Officer and reviewer reports (re-applications only)

19. Table of Contents

20. Proposal

Overview:

Identification of changes in protein-protein interactions associated with sequence variants such as mutations identified as cancer drivers is critical to understanding the consequences of such variation. Technologically speaking, this requires a robust, sensitive and accurate quantification method: in our preliminary work, we have successfully coupled affinity purification (AP) to data-independent mass spectrometric acquisition (SWATH) to analyze the interactome changes imparted by mutations of the kinase CDK4. Here we propose to develop an integrated computational and experimental pipeline that will be scalable to make use of the flood of sequencing data. This project is possible and timely, based on our preliminary data, the availability of disease-associated mutant collections, a strong expertise in structural analysis of protein-protein interactions, and the collaborations we have established with computational biologists and a mass spectrometry vendor. We believe that the AP-SWATH pipeline we are proposing will greatly accelerate biological knowledge regarding the consequences of cancer-associated mutations.

Introduction

Cancer-associated mutations: how interactomics can help reveal function

In recent years, DNA sequencing of healthy and cancerous tissues has revealed a large number of genetic variations¹. With the current speed of next generation sequencing, new mutations are identified daily; yet, there is a dearth of data regarding the consequences of these mutations on cellular functions. While approaches such as RNA interference can help to identify the consequences of depletion of a complete protein, the consequences of sequence variations are more difficult to assess. At the molecular level, in addition to modifying the enzymatic activity of a protein or its expression level, there are many potential outcomes of sequence variation²⁻⁵. One demonstrated consequence of sequence variation is a change in the interactions of the protein, which can contribute to the disease phenotype⁶ (Fig.1). Systematic assessment of protein-protein interaction disruption by yeast two hybrid (Y2H) technologies has revealed clear molecular signatures for certain mutants^{7,8}. However, there are limitations to Y2H, including the fact that it does not easily capture quantitative differences (they are scored as binding / no binding) and that the method in its current implementation is best at highlighting loss rather than gain of interactions.

Affinity purification coupled to mass spectrometry (AP-MS)

AP-MS has been used for many years for the identification of protein-protein interactions⁹ and is attractive since it can identify interactions in near-physiological conditions, providing a context in which a protein

functions. While many groups (including ours¹⁰⁻²⁹) have employed AP-MS to identify static interactomes, limited publications have focused on the identification of differential interactions (reviewed in³⁰). In general, these studies have made use of quantitative proteomics approaches to discriminate between condition-specific interactions. Such quantitative studies have used various techniques, such as spectral counting (e.g.,³¹⁻³⁴), which is limited to monitoring relatively large changes in quantity for relatively abundant proteins, or the intensity of the ions in the precursor ion scans (MS1 or survey scan)³⁵, whose measurements are often problematic. Alternatively, isotopic labeling has been employed for differential proteomics, and this has proven very useful, but fairly expensive and difficult to scale to large numbers of samples^{36,37}. Notably, all of these approaches require MS acquisition in a data-dependent manner (DDA), where the peptides to be sequenced are selected based on their relative abundance in the mass spectrometry MS1 survey scans. DDA introduces stochasticity in the process, which makes it difficult to make absolute conclusions in the absence of a peptide or protein across different samples, especially for lower abundance species³⁸. This is problematic for comparative quantification as is required for mapping sequence variant interactomes³⁹.

Towards better and faster MS measurements

In recent years, a different paradigm for MS-based quantification of proteins has gained increased acceptance. Quantification at the MS/MS (MS2) level increases specificity and signal-to-noise ratios as compared to MS1⁴⁰ and permits robust and sensitive quantification without requiring labeling^{41,42}. Standard MS2 quantification (known as selected reaction monitoring, SRM) however requires a fairly significant investment in assay development for each peptide and/or protein of interest, limiting its usefulness for systems biology experiments. Furthermore, the list of peptide species to be monitored must be predetermined, making it difficult to monitor the gain of interactions that would be generated by a mutation. Recently, a paradigm-changing MS acquisition method has been introduced that has the quantitative benefits of MS2 quantification without the limitations of the classical approaches. In this method, named data-independent acquisition (DIA)^{43,44}, all precursor ions are fragmented independently of their signal in MS1 (Figs.2,3). A type of DIA that is particularly promising for the analysis of AP samples is termed SWATH (Sequential Window Acquisition of all Theoretical spectra)⁴⁵. In SWATH, the entire useful mass range is scanned over a ~ 3 second cycle with windows of incremental mass (here, 25 Da) and all precursors in each window are fragmented, resulting in a complete MS2 map of all compounds that are present in the sample. This map can then be used to detect and quantify proteins in the sample.

Here, we will systematically employ AP-SWATH to study cancer-modulated interactions. We will refine and disseminate an efficient pipeline that will accelerate the study of cancer mechanisms and partially bridge the gap between the genomics efforts and the biology of cancer.

Preliminary results: AP-SWATH analysis of CDK4 sequence variants

In the past years, the Gingras lab has developed a robust experimental pipeline that exploits purification from a single epitope tag (often 3xFLAG), using proteins expressed at near endogenous levels to identify interaction partners^{9,11,13-18,20,22,23,27-29,46-50}. We have further created a database system called ProHits to manage all the MS interaction data^{51,52}, and developed (with A Nesvizhskii) scoring tools to identify true interactors using semi-quantitative approaches^{24,53,54}. To develop the AP-SWATH quantitative proteomics approach, we used two melanoma-associated sequence variants for the cyclin-dependent kinase CDK4, R24C and R24H^{55,56} (Fig.4). These preclude association of CDK4 with a family of polypeptide inhibitors, the INK proteins⁵⁷, resulting in unrepresed CDK4 activity and accelerated cellular proliferation^{7,58}. Consistent with the critical role of the CDK4-INK relationship in cancer, INK proteins are recurrently down-regulated in melanoma and other cancers⁵⁹⁻⁶¹. Quantification of SWATH data was accomplished by a targeted data extraction strategy using a spectral library⁴⁵ which we built from the same

samples by standard DDA. Spectra within this library were used to extract quantitative data from each SWATH run (Fig.3). The extracted data was subjected to statistical analysis⁴¹ using a strategy which evaluates the quality of each peak in MS2 and performs fold change calculation of pairwise samples at the level of MS2 peaks, peptides, and proteins. As expected, both R24C and R24H mutants largely lost interactions with members of the INK inhibitors (Fig.4). Interestingly, the CDK4 mutants also exhibited a markedly increased association with HSP90 proteins, the kinase co-chaperone CDC37, and several other components of the HSP90 network. We validated these regulated interactions by IP-western analysis, in all cases confirming the trends observed by AP-SWATH.

While an increased interaction between CDK4 mutants and HSP90 has never been reported, mutations in several other kinases lead to an increased interaction with CDC37 and HSP90^{62,63}. A model of addiction to HSP90 has been proposed for these mutant kinases that become dependent on CDC37-HSP90 folding for stability and activity⁶⁴. Consequently, disruption of the kinase-HSP90 interaction with HSP90 inhibitors has been proposed as a therapeutic avenue for cancers driven by oncogenic kinases (Fig.5). Using AP-SWATH, we next assessed the consequences on CDK4-HSP90 interactions of treatment with NVP-AUY922 (a potent HSP90 inhibitor currently undergoing clinical trial⁶⁵). As expected, treatment of the cells expressing wild type CDK4 with the HSP90 inhibitor resulted in marked dissociation of CDC37, HSP90 and several – but not all – other proteins by AP-SWATH and IP-western. Importantly, however, and as opposed to the report that other mutant kinases show increased sensitivity for HSP90 inhibitors, we found that the R24C and R24H mutants still significantly bound to CDC37 and HSP90 in the presence of the inhibitor, as compared to the wt; this was again confirmed by IP-Western (Fig.6), and by performing time course and dose-dependency studies (not shown). These results indicate that tumors driven by these CDK4 oncogenic kinases may not readily benefit from treatment with HSP90 inhibitors. Furthermore, our experiments set the stage for analyzing other drug-regulated interactions, including for other oncogenic kinase variants.

Research objectives

Our preliminary data highlights a quantitative pipeline to study interactome changes in an unbiased manner. Here we propose to further extend this study and define what types of mutations are best analyzed with this approach.

Aim 1. Expand the pilot project

Through a collaboration with the Vidal and Hill lab (see letter) we have access to a collection of ~4000 disease-related alleles, many of which are implicated in cancer. In a first step towards selecting additional baits for AP-SWATH analysis, we will refine the list of candidate allelic variants by a) restricting the list to mutants in which the link to cancer is clear (e.g., which are clearly cancer driver mutations); b) eliminating proteins that are localized in membranous components (these are problematic by AP-MS); c) selecting protein-protein interaction pairs for which a mutation is likely to disrupt an interaction (see details below); d) prioritizing proteins expressed endogenously in the model line used for AP-MS^{66,67}; e) from these, priority will be given to baits for which we already have reagents, including antibodies to the endogenous proteins.

The Vidal/Hill collection has already been enriched for their likelihood at altering interactions, but a systematic structural analysis of the possible impact of these mutants on altering interaction surfaces has not been performed so far. The Kim lab, which has ample expertise in this field and indeed pioneered some of these approaches⁶⁸⁻⁷⁰, will build a structure-based analysis pipeline to analyze these different mutants. We have already built a structure-based interaction network using available interactome data and high-quality structural data of known protein interactions (in databases such as the PDB⁷¹, 3DID⁷² and iPfam⁷³). Using this resource we will score all mutations by mapping them to the structure of the proteins and determining their

interaction potential (based on position in the structure and disruption potential of the mutation^{68,69}). Many of the mutations that do not map onto our structure-based interaction network may lie in disordered regions and affect domain-peptide interactions. Such interactions tend to be of particular importance in cancer-related signaling pathways and are another focus of the Kim lab^{74,75}. We will hence also be able to evaluate their effect on the interactions mediated by them.

Using this prioritization and the criteria outlined above, we will define a set of 40 wt baits and a total of ~60 allelic variants (~100 independent cell lines). These will be processed by AP-SWATH; validation will be by AP-SRM⁴¹ and/or by IP-western (as in Fig.4) to yield quantitative information about changes in interactions between the variants.

Aim 2. Study phosphatase mutant interactomes

The Gingras lab has a long-standing interest in the analysis of PP2A-like serine/threonine phosphatases, and we have recently expanded our interaction studies to all 150 human phosphatases and ~25 regulatory subunits for which we have performed AP-MS on stable cell lines expressing all WT proteins (we are using the same cell lines for other functional genomics / proteomics studies). While the scaffolding subunit of PP2A (*PPP2R1A/PPP2R1B*) has long been known to be mutated in cancers^{76,77}, recent reports have accumulated that implicate other phosphatases in cancer^{76,78-80}. Notably, PPP6C was identified as a driver in melanoma in two recent studies^{78,81}. Though many of these mutations are expected to abrogate phosphatase activity, some of the other mutations may regulate interactions. Since we have high density interaction maps for both PPP6C and the PP2A proteins (and antibodies to many components of the network), we will reconstitute each of the point mutations in the Flp-In T-REx system and perform a) phosphatase activity measurements (as in^{29,82}); b) interactome mapping by AP-SWATH; c) further study in relevant cancer cells (for example, in the case of PPP6C, we have raised efficient polyclonal antibodies which can be used for AP-SWATH from endogenous material, in this case, melanoma cells). This activity will be expanded during the timeframe of the grant to ~50 mutant phosphatase alleles.

Aim 3. Expand to drug-regulated interactions

As interactors (especially regulated binders) are identified in Aims 1 and 2, they will be screened against pharmacological compound databases to see whether they are the targets of any drugs. If this is the case, the drugs will be tested for their effect(s) on the interactions, as we have done in Fig.6 for the HSP90 drugs. This will be done by first screening for possible drug hits (the Kim lab has related projects to do just this, using a number of accessible web tools, including DrugBank⁸³ and Therapeutic Target Database, TTD⁸⁴). For initial test screening, compounds can be cherry picked from the large collections available at the Lunenfeld SMART facility (see letter from J Dennis), and larger amounts will be purchased as needed. This last aim should help us accelerate the process of finding drugs that modulate interactions, enabling us to propose steps towards therapeutic modulation of protein-protein interactions, thereby expanding the arsenal of anti-cancer agents.

Perspectives

As we are working towards the completion of Aims 1 and 2, we will be continuously revisiting our experimental and computational pipeline. We are working with mass spectrometry vendor AB Sciex, SLRI colleague T Pawson and computational proteomics expert A Nesvizhskii to refine the SWATH approach and improve the MS data handling. Our proposal benefits from the complementarity in expertise of the lead PIs, our collaborators and our lab members (PDF Couzens and student Veri will perform the experimental objectives assisted by several other Gingras lab members while PDF Teyra will be tasked with the computational predictions of regulated interactions).

We (Kim) will also reanalyze the data generated in Aims 1 and 2 to improve our ability to predict interaction modulation by mutation. Since there is currently a dearth of high quality quantitative interaction proteomics data to analyze the consequences of a mutation on interactome changes, our dataset will enable us to revise the scoring approach and provide an improved stratification strategy for future studies, both within the context of model systems, and with cell lines, xenografts and tissues directly derived from the patient samples.

Taken together, our project will develop an efficient platform to enable higher throughput analysis and provide cancer researchers with annotation of interactome changes imparted by mutations, and when possible, by treatment with pharmaceuticals. This should greatly accelerate the knowledge of the biology of the mutations, and provide the research community with leads for biological follow-up, thereby accelerating the development of therapeutic options. For example, the data generated here will provide a number of highly valuable targets that will be further studied in a related project from the Kim lab that aims at generating peptide-based inhibitors to cancer-related protein-protein interactions using a combination of computational and combinatorial chemistry approaches. By focusing on the interactions that are specifically mutated (e.g., strengthened) in cancer cells and whose inhibition would be of particular value therapeutically, target prioritization for inhibitor development would be improved.

21. Vision statement

What we envision as main outcomes of the project are a) the generation of valuable data on the interactions modulated across ~150 protein forms; b) the creation of a robust pipeline which enables us to combine structural information and experimental data to systematically map the changes in interactomes imparted by cancer associated mutations. As there is an increasing push towards DNA sequencing to identify new mutations, it will be essential to develop methodologies that enable to zoom in on mutations that can affect interactions and to define which of these are regulated interactions. While some of the interactions may be themselves currently “druggable”, as in the case of the HSP90 inhibitors used in our preliminary study, others may become attractive targets for future compound development. We therefore see our studies as helping to bridge the gap between the genomics field and the bedside.

22. Tables, graphs, charts and associated legends

Attachments:

[gingras_anneclaude-figures](#)

23. List of references

- [1] Lander, E. S., Initial impact of the sequencing of the human genome. *Nature* 2011, 470, 187-197.
- [2] Torkamani, A., Verkhivker, G., Schork, N. J., Cancer driver mutations in protein kinase genes. *Cancer Lett* 2009, 281, 117-127.
- [3] Steward, R. E., MacArthur, M. W., Laskowski, R. A., Thornton, J. M., Molecular basis of inherited diseases: a structural perspective. *Trends Genet* 2003, 19, 505-513.
- [4] Lahiry, P., Torkamani, A., Schork, N. J., Hegele, R. A., Kinase mutations in human disease: interpreting genotype-phenotype relationships. *Nat Rev Genet* 2010, 11, 60-74.
- [5] Vacic, V., Iakoucheva, L. M., Disease mutations in disordered regions--exception to the rule? *Mol Biosyst* 2012, 8, 27-32.
- [6] Schuster-Bockler, B., Bateman, A., Protein interactions in human genetic diseases. *Genome Biol* 2008, 9, R9.

- [7] Zhong, Q., Simonis, N., Li, Q. R., Charlotteaux, B., Heuze, F., Klitgord, N., Tam, S., Yu, H., Venkatesan, K., Mou, D., Swearingen, V., Yildirim, M. A., Yan, H., Dricot, A., Szeto, D., Lin, C., Hao, T., Fan, C., Milstein, S., Dupuy, D., Brasseur, R., Hill, D. E., Cusick, M. E., Vidal, M., Edgetic perturbation models of human inherited disorders. *Mol Syst Biol* 2009, 5, 321.
- [8] Dreze, M., Charlotteaux, B., Milstein, S., Vidalain, P. O., Yildirim, M. A., Zhong, Q., Svrzikapa, N., Romero, V., Laloux, G., Brasseur, R., Vandenhaute, J., Boxem, M., Cusick, M. E., Hill, D. E., Vidal, M., 'Edgetic' perturbation of a *C. elegans* BCL2 ortholog. *Nat Methods* 2009, 6, 843-849.
- [9] Gingras, A. C.*, Gstaiger, M., Raught, B., Aebersold, R., Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* 2007, 8, 645-654.
- [10] Ryan, O., Shapiro, R. S., Kurat, C. F., Mayhew, D., Baryshnikova, A., Chin, B., Lin, Z. Y., Cox, M. J., Vizeacoumar, F., Cheung, D., Bahr, S., Tsui, K., Tebbji, F., Sellam, A., Istel, F., Schwarzmuller, T., Reynolds, T. B., Kuchler, K., Gifford, D. K., Whiteway, M., Giaever, G., Nislow, C., Costanzo, M., Gingras, A. C., Mitra, R. D., Andrews, B., Fink, G. R., Cowen, L. E., Boone, C., Global gene deletion analysis exploring yeast filamentous growth. *Science* 2012, 337, 1353-1356.
- [11] Gomez-Ferreria, M. A., Bashkurov, M., Mullin, M., Gingras, A. C., Pelletier, L., CEP192 interacts physically and functionally with the K63-deubiquitinase CYLD to promote mitotic spindle assembly. *Cell Cycle* 2012, 11.
- [12] Belozerov, V. E., Lin, Z. Y., Gingras, A. C., McDermott, J. C., Michael Siu, K. W., High-Resolution Protein Interaction Map of the *Drosophila melanogaster* p38 Mitogen-Activated Protein Kinases Reveals Limited Functional Redundancy. *Mol Cell Biol* 2012, 32, 3695-3706.
- [13] van Zuylen, W. J., Doyon, P., Clement, J. F., Khan, K. A., D'Ambrosio, L. M., Do, F., St-Amant-Verret, M., Wissanji, T., Emery, G., Gingras, A. C., Meloche, S., Servant, M. J., Proteomic profiling of the TRAF3 interactome network reveals a new role for the ER-to-Golgi transport compartments in innate immunity. *PLoS Pathog* 2012, 8, e1002747.
- [14] Costa, B., Kean, M. J., Ast, V., Knight, J. D., Mett, A., Levy, Z., Ceccarelli, D. F., Badillo, B. G., Eils, R., Konig, R., Gingras, A. C., Fainzilber, M., STK25 Protein Mediates TrkA and CCM2 Protein-dependent Death in Pediatric Tumor Cells of Neural Origin. *J Biol Chem* 2012, 287, 29285-29289.
- [15] Morita, M., Ler, L. W., Fabian, M. R., Siddiqui, N., Mullin, M., Henderson, V. C., Alain, T., Fonseca, B. D., Karashchuk, G., Bennett, C. F., Kabuta, T., Higashi, S., Larsson, O., Topisirovic, I., Smith, R. J., Gingras, A. C., Sonenberg, N., A Novel 4EHP-GIGYF2 Translational Repressor Complex Is Essential for Mammalian Development. *Mol Cell Biol* 2012, 32, 3585-3593.
- [16] Jovic, M., Kean, M. J., Szentpetery, Z., Polevoy, G., Gingras, A. C., Brill, J. A., Balla, T., Two phosphatidylinositol 4-kinases control lysosomal delivery of the Gaucher disease enzyme, beta-glucocerebrosidase. *Mol Biol Cell* 2012, 23, 1533-1545.
- [17] Al-Hakim, A. K., Bashkurov, M., Gingras, A. C., Durocher, D., Pelletier, L., Interaction proteomics identify NEURL4 and the HECT E3 ligase HERC2 as novel modulators of centrosome architecture. *Mol Cell Proteomics* 2012, 11, M111 014233.
- [18] Kean, M. J., Ceccarelli, D. F., Goudreault, M., Sanches, M., Tate, S., Larsen, B., Gibson, L. C., Derry, W. B., Scott, I. C., Pelletier, L., Baillie, G. S., Sicheri, F., Gingras, A. C.*, Structure-function analysis of core STRIPAK Proteins: a signaling complex implicated in Golgi polarization. *J Biol Chem* 2011, 286, 25065-25075.
- [19] Li, Z., Vizeacoumar, F. J., Bahr, S., Li, J., Warringer, J., Vizeacoumar, F. S., Min, R., Vandersluis, B., Bellay, J., Devit, M., Fleming, J. A., Stephens, A., Haase, J., Lin, Z. Y., Baryshnikova, A., Lu, H., Yan, Z., Jin, K., Barker, S., Datti, A., Giaever, G., Nislow, C., Bulawa, C., Myers, C. L., Costanzo, M., Gingras, A. C., Zhang, Z., Blomberg, A., Bloom, K., Andrews, B., Boone, C., Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nat Biotechnol* 2011, 29, 361-367.
- [20] Skarra, D. V., Goudreault, M., Choi, H., Mullin, M., Nesvizhskii, A. I., Gingras, A. C.*, Honkanen, R. E.*,

Label-free quantitative proteomics and SAINT analysis enable interactome mapping for the human Ser/Thr protein phosphatase 5. *Proteomics* 2011, 11, 1508-1516.

[21] Templeton, G. W., Nimick, M., Morrice, N., Campbell, D., Goudreault, M., Gingras, A. C., Takemiy, A., Shimazaki, K., Moorhead, G. B., Identification and characterization of Atl-2, an Arabidopsis homologue of an ancient protein phosphatase 1 (PP1) regulatory subunit. *Biochem J* 2011, 435, 73-83.

[22] O'Donnell, L., Panier, S., Wildenhain, J., Tkach, J. M., Al-Hakim, A., Landry, M. C., Escribano-Diaz, C., Szilard, R. K., Young, J. T., Munro, M., Canny, M. D., Kolas, N. K., Zhang, W., Harding, S. M., Ylanko, J., Mendez, M., Mullin, M., Sun, T., Habermann, B., Datti, A., Bristow, R. G., Gingras, A. C., Tyers, M. D., Brown, G. W., Durocher, D., The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Mol Cell* 2010, 40, 619-631.

[23] Nakada, S., Tai, I., Panier, S., Al-Hakim, A., Iemura, S., Juang, Y. C., O'Donnell, L., Kumakubo, A., Munro, M., Sicheri, F., Gingras, A. C., Natsume, T., Suda, T., Durocher, D., Non-canonical inhibition of DNA damage-dependent ubiquitination by OTUB1. *Nature* 2010, 466, 941-946.

[24] Breitkreutz, A., Choi, H., Sharom, J. R., Boucher, L., Neduva, V., Larsen, B., Lin, Z. Y., Breitkreutz, B. J., Stark, C., Liu, G., Ahn, J., Dewar-Darch, D., Reguly, T., Tang, X., Almeida, R., Qin, Z. S., Pawson, T., Gingras, A. C., Nesvizhskii, A. I., Tyers, M., A global protein kinase and phosphatase interaction network in yeast. *Science* 2010, 328, 1043-1046.

[25] Mak, A. B., Ni, Z., Hewel, J. A., Chen, G. I., Zhong, G., Karamboulas, K., Blakely, K., Smiley, S., Marcon, E., Roudeva, D., Li, J., Olsen, J. B., Wan, C., Punna, T., Isserlin, R., Chetyrkin, S., Gingras, A. C., Emili, A., Greenblatt, J., Moffat, J., A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency. *Mol Cell Proteomics* 2010, 9, 811-823.

[26] Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L., Toufighi, K., Mostafavi, S., Prinz, J., St Onge, R. P., VanderSluis, B., Makhnevych, T., Vizeacoumar, F. J., Alizadeh, S., Bahr, S., Brost, R. L., Chen, Y., Cokol, M., Deshpande, R., Li, Z., Lin, Z. Y., Liang, W., Marback, M., Paw, J., San Luis, B. J., Shuteriqi, E., Tong, A. H., van Dyk, N., Wallace, I. M., Whitney, J. A., Weirauch, M. T., Zhong, G., Zhu, H., Houry, W. A., Brudno, M., Ragibizadeh, S., Papp, B., Pal, C., Roth, F. P., Giaever, G., Nislow, C., Troyanskaya, O. G., Bussey, H., Bader, G. D., Gingras, A. C., Morris, Q. D., Kim, P. M., Kaiser, C. A., Myers, C. L., Andrews, B. J., Boone, C., The genetic landscape of a cell. *Science* 2010, 327, 425-431.

[27] Lawo, S., Bashkurov, M., Mullin, M., Ferreria, M. G., Kittler, R., Habermann, B., Tagliaferro, A., Poser, I., Hutchins, J. R., Hegemann, B., Pinchev, D., Buchholz, F., Peters, J. M., Hyman, A. A., Gingras, A. C., Pelletier, L., HAUS, the 8-subunit human Augmin complex, regulates centrosome and spindle integrity. *Curr Biol* 2009, 19, 816-826.

[28] Goudreault, M., D'Ambrosio, L. M., Kean, M. J., Mullin, M. J., Larsen, B. G., Sanchez, A., Chaudhry, S., Chen, G. I., Sicheri, F., Nesvizhskii, A. I., Aebersold, R., Raught, B., Gingras, A. C., A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. *Mol Cell Proteomics* 2009, 8, 157-171.

[29] Chen, G. I., Tisayakorn, S., Jorgensen, C., D'Ambrosio, L. M., Goudreault, M., Gingras, A. C., PP4R4/KIAA1622 forms a novel stable cytosolic complex with phosphoprotein phosphatase 4. *J Biol Chem* 2008, 283, 29273-29284.

[30] Ideker, T., Krogan, N. J., Differential network biology. *Mol Syst Biol* 2012, 8, 565.

[31] Bouwmeester, T., Bauch, A., Ruffner, H., Angrand, P. O., Bergamini, G., Croughton, K., Cruciat, C., Eberhard, D., Gagneur, J., Ghidelli, S., Hopf, C., Huhse, B., Mangano, R., Michon, A. M., Schirle, M., Schlegl, J., Schwab, M., Stein, M. A., Bauer, A., Casari, G., Drewes, G., Gavin, A. C., Jackson, D. B., Joberty, G., Neubauer, G., Rick, J., Kuster, B., Superti-Furga, G., A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol* 2004, 6, 97-105.

[32] Glatter, T., Schittenhelm, R. B., Rinner, O., Roguska, K., Wepf, A., Junger, M. A., Kohler, K., Jevtov, I.,

- Choi, H., Schmidt, A., Nesvizhskii, A. I., Stocker, H., Hafen, E., Aebersold, R., Gstaiger, M., Modularity and hormone sensitivity of the *Drosophila melanogaster* insulin receptor/target of rapamycin interaction proteome. *Mol Syst Biol* 2011, 7, 547.
- [33] Lee, K. K., Sardiu, M. E., Swanson, S. K., Gilmore, J. M., Torok, M., Grant, P. A., Florens, L., Workman, J. L., Washburn, M. P., Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. *Mol Syst Biol* 2011, 7, 503.
- [34] Ho, L., Ronan, J. L., Wu, J., Staahl, B. T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A. I., Ranish, J., Crabtree, G. R., An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A* 2009, 106, 5181-5186.
- [35] Rinner, O., Mueller, L. N., Hubalek, M., Muller, M., Gstaiger, M., Aebersold, R., An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. *Nat Biotechnol* 2007, 25, 345-352.
- [36] Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M., Foster, L. J., Mann, M., A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat Biotechnol* 2003, 21, 315-318.
- [37] Trinkle-Mulcahy, L., Andersen, J., Lam, Y. W., Moorhead, G., Mann, M., Lamond, A. I., Repo-Man recruits PP1 gamma to chromatin and is essential for cell viability. *J Cell Biol* 2006, 172, 679-692.
- [38] Tabb, D. L., Vega-Montoto, L., Rudnick, P. A., Variyath, A. M., Ham, A. J., Bunk, D. M., Kilpatrick, L. E., Billheimer, D. D., Blackman, R. K., Cardasis, H. L., Carr, S. A., Clouser, K. R., Jaffe, J. D., Kowalski, K. A., Neubert, T. A., Regnier, F. E., Schilling, B., Tegeler, T. J., Wang, M., Wang, P., Whiteaker, J. R., Zimmerman, L. J., Fisher, S. J., Gibson, B. W., Kinsinger, C. R., Mesri, M., Rodriguez, H., Stein, S. E., Tempst, P., Paulovich, A. G., Liebler, D. C., Spiegelman, C., Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* 2010, 9, 761-776.
- [39] Gingras, A. C.*, Raught, B., Beyond hairballs: The use of quantitative mass spectrometry data to understand protein-protein interactions. *FEBS Lett* 2012, 586, 2723-2731.
- [40] Picotti, P., Aebersold, R., Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods* 2012, 9, 555-566.
- [41] Bisson, N., James, D. A., Ivosev, G., Tate, S. A., Bonner, R., Taylor, L., Pawson, T., Selected reaction monitoring mass spectrometry reveals the dynamics of signaling through the GRB2 adaptor. *Nat Biotechnol* 2011, 29, 653-658.
- [42] Zheng, Y., Zhang, C., Soliman, M. A., Croucher, D. R., Bagshaw, R., Pascalescu, A., St-Denis, N., Taylor, L., Tate, S., Hardy, R. W., Colwill, K., Dennis, J. W., Dai, A. Y., Daly, R. J., Gingras, A. C., Pawson, T., Dynamic regulation of EGF signaling networks in time and space by the scaffold protein Shc1. *Under Review at Nature* 2012.
- [43] Venable, J. D., Dong, M. Q., Wohlschlegel, J., Dillin, A., Yates, J. R., Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Methods* 2004, 1, 39-45.
- [44] Tate, S., Larsen, B., Bonner, R., Gingras, A. C., Quantitative proteomics trends for protein-protein interactions. *The Journal of Proteomics* 2012, Accepted with minor modifications.
- [45] Gillet, L. C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., Bonner, R., Aebersold, R., Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 2012, 11, O111 016717.
- [46] Kean, M. J., Couzens, A. L., Gingras, A. C.*, Mass spectrometry approaches to study mammalian kinase and phosphatase associated proteins. *Methods* 2012, 57, 400-408.
- [47] Dunham, W. H., Mullin, M., Gingras, A. C.*, Affinity-purification coupled to mass spectrometry: basic principles and strategies. *Proteomics* 2012, 12, 1576-1590.
- [48] Dunham, W. H., Larsen, B., Tate, S., Badillo, B. G., Goudreault, M., Tehami, Y., Kislinger, T., Gingras, A. C.*, A cost-benefit analysis of multidimensional fractionation of affinity purification-mass spectrometry samples. *Proteomics* 2011, 11, 2603-2612.

- [49] Chen, G. I., Gingras, A. C.*, Affinity-purification mass spectrometry (AP-MS) of serine/threonine phosphatases. *Methods* 2007, 42, 298-305.
- [50] Gingras, A. C.*, Caballero, M., Zarske, M., Sanchez, A., Hazbun, T. R., Fields, S., Sonenberg, N., Hafen, E., Raught, B., Aebersold, R., A novel, evolutionarily conserved protein phosphatase complex involved in cisplatin sensitivity. *Mol Cell Proteomics* 2005, 4, 1725-1740.
- [51] Liu, G., Zhang, J., Choi, H., Lambert, J. P., Srikumar, T., Larsen, B., Nesvizhskii, A. I., Raught, B., Tyers, M., Gingras, A. C.*, Using ProHits to Store, Annotate, and Analyze Affinity Purification-Mass Spectrometry (AP-MS) Data. *Curr Protoc Bioinformatics* 2012, Chapter 8, Unit8 16.
- [52] Liu, G., Zhang, J., Larsen, B., Stark, C., Breitkreutz, A., Lin, Z. Y., Breitkreutz, B. J., Ding, Y., Colwill, K., Pascalescu, A., Pawson, T., Wrana, J. L., Nesvizhskii, A. I., Raught, B., Tyers, M.* , Gingras, A. C.*, ProHits: integrated software for mass spectrometry-based interaction proteomics. *Nat Biotechnol* 2010, 28, 1015-1017.
- [53] Choi, H., Liu, G., Mellacheruvu, D., Tyers, M., Gingras, A. C.*, Nesvizhskii, A. I.*, Analyzing Protein-Protein Interactions from Affinity Purification-Mass Spectrometry Data with SAINT. *Curr Protoc Bioinformatics* 2012, Chapter 8, Unit8 15.
- [54] Choi, H., Larsen, B., Lin, Z. Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z. S., Tyers, M., Gingras, A. C.*, Nesvizhskii, A. I.*, SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat Methods* 2011, 8, 70-73.
- [55] Wolfel, T., Hauer, M., Schneider, J., Serrano, M., Wolfel, C., Klehmann-Hieb, E., De Plaein, E., Hankeln, T., Meyer zum Buschenfelde, K. H., Beach, D., A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma. *Science* 1995, 269, 1281-1284.
- [56] Zuo, L., Weger, J., Yang, Q., Goldstein, A. M., Tucker, M. A., Walker, G. J., Hayward, N., Dracopoli, N. C., Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. *Nat Genet* 1996, 12, 97-99.
- [57] Serrano, M., Hannon, G. J., Beach, D., A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature* 1993, 366, 704-707.
- [58] Coleman, K. G., Wautlet, B. S., Morrissey, D., Mulheron, J., Sedman, S. A., Brinkley, P., Price, S., Webster, K. R., Identification of CDK4 sequences involved in cyclin D1 and p16 binding. *J Biol Chem* 1997, 272, 18869-18874.
- [59] Serrano, M., Lee, H., Chin, L., Cordon-Cardo, C., Beach, D., DePinho, R. A., Role of the INK4a locus in tumor suppression and cell mortality. *Cell* 1996, 85, 27-37.
- [60] Nobori, T., Miura, K., Wu, D. J., Lois, A., Takabayashi, K., Carson, D. A., Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. *Nature* 1994, 368, 753-756.
- [61] Kamb, A., Gruis, N. A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S. V., Stockert, E., Day, R. S., 3rd, Johnson, B. E., Skolnick, M. H., A cell cycle regulator potentially involved in genesis of many tumor types. *Science* 1994, 264, 436-440.
- [62] Shimamura, T., Lowell, A. M., Engelman, J. A., Shapiro, G. I., Epidermal growth factor receptors harboring kinase domain mutations associate with the heat shock protein 90 chaperone and are destabilized following exposure to geldanamycins. *Cancer Res* 2005, 65, 6401-6408.
- [63] Grbovic, O. M., Basso, A. D., Sawai, A., Ye, Q., Friedlander, P., Solit, D., Rosen, N., V600E B-Raf requires the Hsp90 chaperone for stability and is degraded in response to Hsp90 inhibitors. *Proc Natl Acad Sci U S A* 2006, 103, 57-62.
- [64] Prodromou, C., Strategies for stalling malignancy: targeting cancer's addiction to Hsp90. *Curr Top Med Chem* 2009, 9, 1352-1368.
- [65] Brough, P. A., Aherne, W., Barril, X., Borgognoni, J., Boxall, K., Cansfield, J. E., Cheung, K. M., Collins, I., Davies, N. G., Drysdale, M. J., Dymock, B., Eccles, S. A., Finch, H., Fink, A., Hayes, A., Howes, R., Hubbard, R. E., James, K., Jordan, A. M., Lockie, A., Martins, V., Massey, A., Matthews, T. P., McDonald, E., Northfield, C. J., Pearl, L. H., Prodromou, C., Ray, S., Raynaud, F. I., Roughley, S. D., Sharp, S. Y., Surgenor,

- A., Walmsley, D. L., Webb, P., Wood, M., Workman, P., Wright, L., 4,5-diarylisoxazole Hsp90 chaperone inhibitors: potential therapeutic agents for the treatment of cancer. *J Med Chem* 2008, 51, 196-218.
- [66] Desiere, F., Deutsch, E. W., King, N. L., Nesvizhskii, A. I., Mallick, P., Eng, J., Chen, S., Eddes, J., Loevenich, S. N., Aebersold, R., The PeptideAtlas project. *Nucleic Acids Res* 2006, 34, D655-658.
- [67] Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., Wernerus, H., Bjorling, L., Ponten, F., Towards a knowledge-based Human Protein Atlas. *Nat Biotechnol* 2010, 28, 1248-1250.
- [68] Ellis, J. D., Barrios-Rodiles, M., Colak, R., Irimia, M., Kim, T., Calarco, J. A., Wang, X., Pan, Q., O'Hanlon, D., Kim, P. M.*, Wrana, J. L.*., Blencowe, B. J.*., Tissue-specific alternative splicing remodels protein-protein interaction networks. *Mol Cell* 2012, 46, 884-892.
- [69] Kim, P. M., Lu, L. J., Xia, Y., Gerstein, M. B., Relating three-dimensional structures to protein networks provides evolutionary insights. *Science* 2006, 314, 1938-1941.
- [70] Kim, P. M., Korbel, J. O., Gerstein, M. B., Positive selection at the protein network periphery: evaluation in terms of structural constraints and cellular context. *Proc Natl Acad Sci U S A* 2007, 104, 20274-20279.
- [71] Sussman, J. L., Lin, D., Jiang, J., Manning, N. O., Prilusky, J., Ritter, O., Abola, E. E., Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallogr D Biol Crystallogr* 1998, 54, 1078-1084.
- [72] Stein, A., Ceol, A., Aloy, P., 3did: identification and classification of domain-based interactions of known three-dimensional structure. *Nucleic Acids Res* 2011, 39, D718-723.
- [73] Finn, R. D., Marshall, M., Bateman, A., iPfam: visualization of protein-protein interactions in PDB at domain and amino acid resolutions. *Bioinformatics* 2005, 21, 410-412.
- [74] Gfeller, D., Butty, F., Wierzbicka, M., Verschueren, E., Vanhee, P., Huang, H., Ernst, A., Dar, N., Stagljar, I., Serrano, L., Sidhu, S. S., Bader, G. D., Kim, P. M.*, The multiple-specificity landscape of modular peptide recognition domains. *Mol Syst Biol* 2011, 7, 484.
- [75] Bellay, J., Han, S., Michaut, M., Kim, T., Costanzo, M., Andrews, B. J., Boone, C., Bader, G. D., Myers, C. L., Kim, P. M.*, Bringing order to protein disorder through comparative genomics and genetic interactions. *Genome Biol* 2011, 12, R14.
- [76] Jones, S., Wang, T. L., Shih le, M., Mao, T. L., Nakayama, K., Roden, R., Glas, R., Slamon, D., Diaz, L. A., Jr., Vogelstein, B., Kinzler, K. W., Velculescu, V. E., Papadopoulos, N., Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science* 2010, 330, 228-231.
- [77] Wang, S. S., Esplin, E. D., Li, J. L., Huang, L., Gazdar, A., Minna, J., Evans, G. A., Alterations of the PPP2R1B gene in human lung and colon cancer. *Science* 1998, 282, 284-287.
- [78] Krauthammer, M., Kong, Y., Ha, B. H., Evans, P., Bacchiocchi, A., McCusker, J. P., Cheng, E., Davis, M. J., Goh, G., Choi, M., Ariyan, S., Narayan, D., Dutton-Regester, K., Capatana, A., Holman, E. C., Bosenberg, M., Sznol, M., Kluger, H. M., Brash, D. E., Stern, D. F., Materin, M. A., Lo, R. S., Mane, S., Ma, S., Kidd, K. K., Hayward, N. K., Lifton, R. P., Schlessinger, J., Boggon, T. J., Halaban, R., Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nat Genet* 2012, 44, 1006-1014.
- [79] Julien, S. G., Dube, N., Hardy, S., Tremblay, M. L., Inside the human cancer tyrosine phosphatome. *Nat Rev Cancer* 2011, 11, 35-49.
- [80] Jones, D. T., Jager, N., Kool, M., Zichner, T., Hutter, B., Sultan, M., Cho, Y. J., Pugh, T. J., Hovestadt, V., Stutz, A. M., Rausch, T., Warnatz, H. J., Ryzhova, M., Bender, S., Sturm, D., Pleier, S., Cin, H., Pfaff, E., Sieber, L., Wittmann, A., Remke, M., Witt, H., Hutter, S., Tzaridis, T., Weischenfeldt, J., Raeder, B., Avci, M., Amstislavskiy, V., Zapatka, M., Weber, U. D., Wang, Q., Lasitschka, B., Bartholomae, C. C., Schmidt, M., von Kalle, C., Ast, V., Lawerenz, C., Eils, J., Kabbe, R., Benes, V., van Sluis, P., Koster, J., Volckmann, R., Shih, D., Betts, M. J., Russell, R. B., Coco, S., Tonini, G. P., Schuller, U., Hans, V., Graf, N., Kim, Y. J., Monoranu, C., Roggendorf, W., Unterberg, A., Herold-Mende, C., Milde, T., Kulozik, A. E., von Deimling, A., Witt, O., Maass, E., Rossler, J., Ebinger, M., Schuhmann, M. U., Fruhwald, M. C., Hasselblatt, M., Jabado, N.,

Rutkowski, S., von Bueren, A. O., Williamson, D., Clifford, S. C., McCabe, M. G., Collins, V. P., Wolf, S., Wiemann, S., Lehrach, H., Brors, B., Scheurlen, W., Felsberg, J., Reifenberger, G., Northcott, P. A., Taylor, M. D., Meyerson, M., Pomeroy, S. L., Yaspo, M. L., Korbel, J. O., Korshunov, A., Eils, R., Pfister, S. M., Lichter, P., Dissecting the genomic complexity underlying medulloblastoma. *Nature* 2012, 488, 100-105.

[81] Hodis, E., Watson, I. R., Kryukov, G. V., Arold, S. T., Imielinski, M., Theurillat, J. P., Nickerson, E., Auclair, D., Li, L., Place, C., Dicara, D., Ramos, A. H., Lawrence, M. S., Cibulskis, K., Sivachenko, A., Voet, D., Saksena, G., Stransky, N., Onofrio, R. C., Winckler, W., Ardlie, K., Wagle, N., Wargo, J., Chong, K., Morton, D. L., Stemke-Hale, K., Chen, G., Noble, M., Meyerson, M., Ladbury, J. E., Davies, M. A., Gershenwald, J. E., Wagner, S. N., Hoon, D. S., Schadendorf, D., Lander, E. S., Gabriel, S. B., Getz, G., Garraway, L. A., Chin, L., A landscape of driver mutations in melanoma. *Cell* 2012, 150, 251-263.

[82] Nakada, S., Chen, G. I., Gingras, A. C.*, Durocher, D.*, PP4 is a gamma H2AX phosphatase required for recovery from the DNA damage checkpoint. *EMBO Rep* 2008, 9, 1019-1026.

[83] Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., Wishart, D. S., DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res* 2011, 39, D1035-1041.

[84] Zhu, F., Han, B., Kumar, P., Liu, X., Ma, X., Wei, X., Huang, L., Guo, Y., Han, L., Zheng, C., Chen, Y., Update of TTD: Therapeutic Target Database. *Nucleic Acids Res* 2010, 38, D787-791.

24. Appendices

25. Disclosure of commercial interest related to this application

- a. All applicants confirm that they have no commercial interest to declare (Please select Yes No
Yes or No)
- b. If no, provide description of commercial Interests

Budget

Description	2013	2014	2015	Total
DIRECT EXPENSES				
Program Expenses				
1 Supplies and Expenses				gingras_anneclaude-justification-supplies.pdf
Mass spectrometry consumables	7,500.00	7,500.00	0.00	15,000.00
Cell culture consumables	6,000.00	6,000.00	0.00	12,000.00
Affinity reagents consumables	6,000.00	6,000.00	0.00	12,000.00
Molecular biology / general consumables	5,200.00	5,200.00	0.00	10,400.00
Publication fees	1,500.00	1,500.00	0.00	3,000.00
Travel	1,500.00	1,500.00	0.00	3,000.00
Total for Supplies and Expenses	27,700.00	27,700.00	0.00	55,400.00
2 Salaries and Wages				gingras_anneclaude-justification_salaries.pdf
Amber Couzens, PDF, year 3, 100% effort [Amber Couzens, Postdoctoral Fellow, year 3, 100% effort, 100% salary support requested from CCSRI Supervised by Anne-Claude Gingras]	38,165.00	38,165.00	0.00	76,330.00
Joan Teyra, PDF, year 2, 40% effort [Joan Teyra, Postdoctoral Fellow, year 2, 40% effort, 40% salary support requested from CCSRI Supervised by Philip Kim]	14,632.00	14,632.00	0.00	29,264.00
Amanda Veri, PhD student, 100% effort [Amanda Veri, Graduate Student, 100% effort, 100% salary support requested from CCSRI Supervised by Anne-Claude Gingras]	19,500.00	19,500.00	0.00	39,000.00

Budget request for Innovation Grants 1 - 2013
 Applicant: Anne-Claude Gingras
 Application : Rapid mapping of cancer interactomes

10/1/2012

Description	2013	2014	2015	Total
Total for Salaries and Wages	72,297.00	72,297.00	0.00	144,594.00
Total Program Expenses	99,997.00	99,997.00	0.00	199,994.00
Equipment				
1 Permanent Equipment				
Total Equipment	0.00	0.00	0.00	0.00
TOTAL DIRECT EXPENSES	99,997.00	99,997.00	0.00	199,994.00
TOTAL EXPENDITURES	99,997.00	99,997.00	0.00	199,994.00

Budget summary for Innovation Grants 1 - 2013
Applicant: Anne-Claude Gingras
Application : Rapid mapping of cancer interactomes

10/1/2012

Description	2013	2014	2015	Total
Supplies and Expenses	27,700.00	27,700.00	0.00	55,400.00
Salaries and Wages	72,297.00	72,297.00	0.00	144,594.00
Permanent Equipment	0.00	0.00	0.00	0.00
TOTALS	99,997.00	99,997.00	0.00	199,994.00

Other funding

FOR OFFICE USE ONLY:	Version # _____	APP # _____
----------------------	-----------------	-------------

27. Summary of other funding applied for and received

Attachments:

[gingras_anneclaude-other_support](#)

Review panel

FOR OFFICE USE ONLY: Version # _____ APP # _____

28. Panel recommendation

- a. 1st choice of panel suggested by applicant(s): I1a - Biomarkers and Genomics
- b. 2nd choice of panel suggested by applicant(s): I1b - Gene Regulation
- c. Assigned Panel

29. Reviewer recommendation

Name	Department	Institution	Phone no.	E-mail address	Areas of expertise
Michael Washburn	Director of Proteomics Center	Stowers	816.926.4457	mpw@stowers.org	interaction proteomics
Philippe Roux	IRIC	IRIC, U Montreal	514 343.6399	philippe.roux@umontreal.ca	signalisation cellulaire et protéomique
Marius Ueffing	research unit protein science	Helmholtz Zentrum München	49(0)89 3187 3567	marius.ueffing@helmholtz-muenchen.de	functional proteomics

30. Reviewer exclusions

Name	Reason for exclusion

Tracking

FOR OFFICE USE ONLY: Version # _____ APP # _____

31. Research tracking information

31.a. Research focus

I. SECTION I - Research focus (select ONE only)

- Biomedical Research Clinical Research
 Health Services/Systems Research Social, Cultural, Environmental and Population Health

31.b. Research subject

II. SECTION II - Research subject (select ONE or MORE) (Please select applicable values)

Patients/Study Population

- Adult Pediatric

Patient Tissue

- Adult Pediatric

Model System

- Mouse Drosophila C. elegans Zebrafish

- Yeast Other

Cell System

- hESC Tumour Initiating Cell Non-embryo-derived stem cells (eg. iPS cells)

31.c. Cancer site relevance

III. SECTION III - Cancer site relevance

Cancer site relevance	Percentage	Details
All Sites	100	

31.d. Common Scientific Outline (CSO)

IV. Section IV - Common Scientific Outline (CSO) (Please select applicable values)

Biology

- 1.1 - Normal functioning
 1.2 - Cancer initiation: alterations in chromosomes
 1.3 - Cancer initiation: oncogenes and tumour suppressor genes
 1.4 - Cancer progression and metastasis
 1.5 - Resources and infrastructure

Etiology

- 2.1 - Exogenous factors in the origin and cause of cancer
 2.2 - Endogenous factors in the origin and cause of cancer
 2.3 - Interactions of genes/genetic polymorphisms with exogenous/endogenous factors
 2.4 - Resources and infrastructure related to etiology

Prevention

- 3.1 - Interventions to prevent cancer: personal behaviours that affect cancer risk
 3.2 - Nutritional science in cancer prevention

- 3.3 - Chemoprevention
- 3.4 - Vaccines
- 3.5 - Complementary and alternative prevention approaches
- 3.6 - Resources and infrastructure related to prevention

Early Detection, Diagnosis and Prognosis

- 4.1 - Technology development and/or marker discovery
- 4.2 - Technology and/or marker evaluation with respect to fundamental parameters of method
- 4.3 - Technology and/or marker testing in a clinical setting
- 4.4 - Resources and infrastructure related to detection, diagnosis and prognosis

Treatment

- 5.1 - Localized therapies – discovery and development
- 5.2 - Localized therapies – clinical applications
- 5.3 - Systemic therapies – discovery and development
- 5.4 - Systemic therapies – clinical applications
- 5.5 - Combinations of localized and systemic therapies
- 5.6 - Complementary and alternative treatment approaches
- 5.7 - Resources and infrastructure related to treatment

Cancer Control, Survivorship and Outcomes Research

- 6.1 - Patient care and survivorship issues
- 6.2 - Surveillance
- 6.3 - Behaviour
- 6.4 - Cost analyses and healthcare delivery
- 6.5 - Education and communication
- 6.6 - End-of-life care
- 6.7 - Ethics and confidentiality in cancer research
- 6.8 - Complementary and alternative approaches for supportive care of patients and survivors
- 6.9 - Resources and infrastructure related to cancer control, survivorship and outcomes research

Release form

FOR OFFICE USE ONLY:

Version # _____

APP # _____

32. Release form

On condition that:

- the specified information will be shared by CCSRI only with potential donors and for the sole purpose of
 - obtaining additional funding for CCSRI's grant competitions.
 - potential donors will be required to declare conflict of interest, and sign a confidentiality agreement before the
 - specified information is released to them by CCSRI.
- it will be held confidential by them and not released to other parties, and will be returned to CCSRI or
 - destroyed if the decision is not to fund.
- all information released may be retained by the potential donors if it decides to fund the application, and may
 - be used by the donor in its funding announcements and other communications.

I consent to the sharing of the information specified (described above) with potential donors. (Please select Yes or No)

Yes No

Head of Department

FOR OFFICE USE ONLY:

Version # _____

APP # _____

33. Head of Department/Dean confirmation

I confirm that I am aware of the contents of the application being submitted.

Yes No

Name of the Head of Department or Dean

James Woodgett

Title

Director of Research

Date

09/26/2012

Executive authority - research host

FOR OFFICE USE ONLY: Version # _____ APP # _____

34. Executive authority of the host research institution

I confirm that I have read and understood the Host Institution / CCS Agreement Yes No and agree to abide by the terms.

Name of the executive authority - research host Joan Sproul

Date 09/28/2012

Executive authority - financial host

FOR OFFICE USE ONLY:	Version # _____	APP # _____
----------------------	-----------------	-------------

35. Executive authority of the host finance institution

I confirm that I have read and understood the Host Institution / CCS Agreement, Yes No and agree to abide by the terms.

Name of the executive authority - financial host

Title

Date

Post submission publications

FOR OFFICE USE ONLY:	Version # _____	APP # _____
----------------------	-----------------	-------------

36. Post submission publications

CCSRI

FOR OFFICE USE ONLY:	Version # _____	APP # _____
----------------------	-----------------	-------------

37. CCSRI revised lay summary

Attachments Index

FOR OFFICE USE ONLY:	Version # _____	APP # _____
----------------------	-----------------	-------------

#	Section	Title	File Name
1	Background	Curriculum Vitae (CV)	Kim_philip-CV.pdf
2	Applicant info	Gingras_anneclaude-CV	Gingras_anneclaude-CV.pdf
3	Applicant info	gingras_anneclaude-collaborations	gingras_anneclaude-collaborations.pdf
4	Proposal	gingras_anneclaude-figures	gingras_anneclaude-figures.pdf
5	Budget	gingras_anneclaude-justification-supplies.pdf	gingras_anneclaude-justification-supplies.pdf
6	Budget	gingras_anneclaude-justification_salaries.pdf	gingras_anneclaude-justification_salaries.pdf
7	Other funding	gingras_anneclaude-other_support	gingras_anneclaude-other_support.pdf

BIOGRAPHICAL SKETCH: Philip M. KIM

A. Name and Affiliation

Dr. Philip M Kim, Assistant Professor, University of Toronto
Address: 160 College St., Rm 606, CCBR, University of Toronto, Toronto, ON M5S 3E1, Canada
Phone: +1-416-946-3419, Email: pm.kim@utoronto.ca

B. Academic and Training Background

Massachusetts Institute of Technology (August 1998 - February 2003)
Ph.D., Artificial Intelligence Laboratory and Program in Physical Chemistry. GPA 4.8/5.0

University of North Carolina at Chapel Hill (September 1997 - May 1998)
Research Exchange Student, Graduate Program in Cellular and Molecular Biophysics

University of Tuebingen (October 1994 - August 1997)
Vordiplom in Physics (German equivalent to B.S.), Perfect GPA
Vordiplom in Biochemistry (German equivalent to B.S.), Perfect GPA

C. Work Experience

University of Toronto, Terrence Donnelly Center for Cellular and Biomolecular Research, Banting and Best Department of Medical Research, Department of Molecular Genetics and Department of Computer Science (from January 2009 - present)
Assistant Professor, Tenure-stream

Yale University, Department of Molecular Biophysics & Biochemistry (December 2004 - November 2008)
Postdoctoral Associate / Associate Research Scientist. Advisor: Professor Mark B. Gerstein

McKinsey & Company (August 2003 - November 2004)
Associate. Management consulting of high-tech and pharmaceutical companies

Teaching Experience

University of Toronto, Department of Molecular Genetics (September 2010-December 2010, September 2011-December 2011)
Created and taught new class: "A practical course in programming for biologists"

University of Toronto, Department of Molecular Genetics (September 2011-May 2012)
Coordinated graduate class: "Student Seminar Series"

Yale University, Department of Molecular Biophysics and Biochemistry (January 2005 – November 2008)
Mentored a total of 8 graduate/rotation students. Designed research projects and supervised their progress

Massachusetts Institute of Technology, Department of Chemistry (September 2001 - January 2002)
Head Teaching Assistant, Class: "5.60 Thermodynamics and Kinetics"

D. Selected Accomplishments and Honours

i) Honors/Awards

Ontario Research Fund Early Researcher Award (\$100,000) 2012
Genome Technology "PI of tomorrow" Award 2012
Connaught Award 2009

Boehringer Ingelheim Fonds Ph.D. Fellow 2001-2003
 Merck Graduate Fellow in Bioinformatics 1999-2001
 Howard Hughes Medical Institute Predoctoral Fellowship (Honorable Mention) 1999
 Recipient of a study abroad fellowship from the University of Tuebingen 1997-1998
 Award for the best Chemistry Student from the German Chemical Industry 1993
 3rd West-German Champion in Judo 1993, Competitor in the State Judo League 1993-1994

ii) Professional Activities

Editorial Board Member of *PLoS ONE*
 Scientific Advisory Board Member of Reflexion Pharmaceuticals
 Scientific Consultant for Synbody Pharmaceuticals
 Review Editorial Board of *Frontiers in Plant Physiology*
 Referee for *PLoS Computational Biology*, *Molecular Systems Biology*, *BMC Bioinformatics*, *Annual Review of Biochemistry*, *Pattern Recognition*, *Bioinformatics*, *Gene*, *Proteins*, *Nucleic Acids Research*, *PLoS ONE*, *Protein Science*, *Proceedings of the National Academy of Science*, *FEBS letters*, *Molecular and Cellular Proteomics*, *Genome Research*, *Molecular BioSystems*, *Nature Methods*, *Genome Biology*
 Grant reviewer for *Canada Foundation for Innovation*, *Israeli Science Foundation*, *Fonds Wetenschappelijk Onderzoek (Flemish Science Foundation)*
 Member of the Program Committee for ECCB 2012, ECCB 2010, ISMB 2009, ISMB 2008, ISMB/ECCB 2007
 Organizer of the PRD Specificity Prediction Challenge, DREAM4 2009

iii) Invited Talks

ECCB 2012 Highlight Talk*, Basel, Switzerland, September, 2012
 University of California at San Francisco, Department of Bioengineering, San Francisco, CA, July 2012
 Biomathematics and Statistics Symposium, University of Guelph, June 2012
 Indiana University Purdue University, Center for Computational Biology and Bioinformatics, March 2012
 Universite Libre Brussels, Department of Computer Science, February 2012
 Gene Regulation Workshop, Holetown, Barbados, January 2012
 GIW / KSSB 2012, Busan, Korea, December 2011
 Yonsei University, Department of Biotechnology, Seoul, Korea, December 2011
 New York University, Center for Genomics and Systems Biology, New York, NY, October 2011
 FEBS Workshop on Modular Protein Domains and Networks in Disease*, Seefeld, Austria, September 2011
 ISMB/ECCB 2011 Highlight Talk*, Vienna, Austria, July 2011
 University of Vienna, Department of Computational Systems Biology, Vienna, Austria, July 2011
 Winter School in Computational Biology, Brisbane, Queensland, Australia, June 2011
 ESF-EMBO Conference on Protein Interactions*, Costa Brava, Spain, November 2010
 Systems Biology and New Sequencing Tech, Center for Genomic Regulation, Barcelona, Spain, May 2010
 Dialogue on Reverse Engineering Assessment and Methods (DREAM 4), Broad Institute of MIT and Harvard, Cambridge, MA, December 2009
 FEBS Workshop on Modular Protein Domains in Disease*, Seefeld, Austria, September 2009
 Genome Biology and Bioinformatics Retreat, Toronto, ON, May 2009
 University of Toronto, Department of Molecular Genetics, Toronto, ON, March 2009
 Korean Conference on Systems Biology and Bioinformatics, Cheongju, Korea, November 2008
 University of Cambridge, Department of Biochemistry, Cambridge, UK, July 2008
 National University of Singapore, Department of Biochemistry, Singapore, May 2008
 University of Toronto, Terrence Donnelly Centre for Cellular and Biomolecular Research, Toronto, ON, May 2008
 University of California Davis, Department of Computer Science, Davis, CA, May 2008

Spanish National Cancer Research Center (CNIO), Madrid, Spain, April 2008
 Scripps Research Institute, San Diego, CA, April 2008
 Fred Hutchinson Cancer Research Center, Seattle, WA, March 2008
 Carnegie-Mellon University, Lane Center for Computational Biology, Pittsburgh, PA, March 2008
 University of California Irvine, Department of Computer Science, Irvine, CA, March 2008
 University of Maryland, Department of Bioengineering, College Park, MD, February 2008
 Universitaet Duesseldorf, Fakultaet fuer Informatik, Duesseldorf, Germany, February 2008
 Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK, February 2008
 EMBL-EBI, Hinxton, Cambridge, UK, February 2008
 Max Planck Institutes Selection Symposium, Berlin, Germany, February 2008
 University of California Davis, Genome Center, Davis, CA, January 2008
 Georgia Institute of Technology, School of Biology, Atlanta, GA, January 2008
 EMBL/CRG, Systems Biology Unit, Barcelona, Spain, January 2008
 University of Pennsylvania, Department of Biology, Philadelphia, PA, January 2008

*denotes refereed talk

E. Publications since 2008

Papers submitted for publication

Vizeacoumar FJ, **Arnold R**, Vizeacoumar FS, Sayad A, Young JTF, Kwan JHM, Baryshnikova A, Pelte N, Lawo S, Tanaka H, Brown KR, Mak AB, Fedyshyn Y, Mero P, Glauber CB, Lu W, Kasimer D, Maknevych T, Ketela T, Datti A, Emili A, Fraser A, Pelletier L, Wrana J, **Kim PM**, Andrews B, Boone C, Moffat J. *Nature Biotech* (submitted). =31.09

Colak R*, **Kim T***, Michaut M, **Sun MFG**, Irimia M, Bellay J, Myers CL, Blencowe BJ, **Kim PM**. Distinct types of disorder in the human proteome: functional implications for alternative splicing. *PLoS Comp Biol* (submitted) =5.515

Publications (in press and published)

Ellis JD*, Barrios-Rodiles M*, **Colak R***, Irimia M, **Kim TH**, Calarco JA, Wang X, Pan Q, O'hanlon D, **Kim PM#**, Wrana JL#, Blencowe BJ#. Tissue-specific alternative splicing remodels protein-protein interaction networks. *Mol Cell* 2012 Jun 29;46(6):884-92. =14.194

Arnold R, Boonen K, Sun MFG, Kim PM. Computational analysis of interactomes: Current and future perspectives for bioinformatics approaches to model the host-pathogen interaction space. *Methods* 2012 Aug;57(4):508-18. =4.527

Teyra J, Sidhu SS, Kim PM. Elucidation of the binding preferences of peptide recognition modules: SH3 and PDZ domains. *FEBS Lett*. 2012 Jun 9. [Epub ahead of print]. =3.601

Hooda Y, Kim PM. Computational structural analysis of protein interactions and networks. *Proteomics*. 2012 May;12(10):1697-705. doi: 10.1002/pmic.201100597. =4.815

Sun MFG*, Sikora M*, Costanzo M, Boone C, **Kim PM**. Network Evolution: Rewiring and Signatures of Conservation in Signaling. *PLoS Comp. Biol.* 2012 Mar;8(3):e100241. =5.515

Bellay J, **Michaut M, Kim TH, Han S, Colak R**, Myers CL, **Kim PM**. An Omics Perspective on Protein Disorder. *Molecular Biosystems* 2012 Jan 1;8(1): 185-93. =3.825

Kim TH, Tyndel M, Bader GD, Gfeller D#, Kim PM#. MUSI: An integrated system to identify multiple specificity in large peptide or nucleotide data sets. *Nucleic Acids Research* 2011 Dec 30. =7.836

Yip KY*, Utz L*, Sitwell S, Turk BE, Gerstein MB#, Kim PM#. Identification of Specificity Determining Residues in Peptide Recognition Domains using an Information Theoretic Approach Applied to Large-Scale Binding Maps. *BMC Biology* 2011, 9:53. =5.203

Sun MFG, Kim PM. Evolution of Biological Networks. *Genome Biology* 2011 Dec 28;12(12):235. =6.885

Gfeller D, Butty F, Wierzbicka M, Verschueren E, Vanhee P, Huang H, Ernst A, Dar N, Stagljar I, Serrano L, Sidhu SS, Bader GD, Kim PM. The multiple-specificity landscape of modular peptide recognition domains. *Mol Syst Biol.* 2011 Apr 26;7:484. =9.667

Bellay J, Atluri G, Sing TL, Toufighi K, Costanzo M, Ribeiro PS, Pandey G, Baller J, Vandersluis B, Michaut M, Han S, Kim P, Brown GW, Andrews BJ, Boone C, Kumar V, Myers CL. Putting genetic interactions in context through a global modular decomposition. *Genome Res.* 2011 Aug;21(8):1375-87. =13.588

Fasolo J, Sboner A, Sun MG, Yu H, Chen R, Sharon D, Kim PM, Gerstein M, Snyder M. Diverse protein kinase interactions identified by protein microarrays reveal novel connections between cellular processes. *Genes Dev.* 2011 Apr 1;25(7):767-78. =12.889

Bellay J, Han S, Michaut M, Kim T, Costanzo M, Andrews BJ, Boone C, Bader GD, Myers CL, Kim PM. Bringing order to protein disorder through comparative genomics and genetic interactions. *Genome Biol.* 2011 Feb 16;12(2):R14. = 6.885

Shou C, Bhardwaj N, Lam HY, Yan KK, Kim PM, Snyder M, Gerstein MB. Measuring the evolutionary rewiring of biological networks. *PLoS Comput Biol.* 2011 Jan 6;7(1):e1001050. =5.515

Bhardwaj N, Kim PM#, Gerstein MB#. Rewiring of transcriptional regulatory networks: hierarchy, rather than connectivity, better reflects the importance of regulators. *Sci Signal.* 2010 Nov 2;3(146):ra79. [# joint corresponding author]. =6.354

Ernst A, Gfeller D, Kan Z, Seshagiri S, Kim PM, Bader GD, Sidhu SS. Coevolution of PDZ domain-ligand interactions analyzed by high-throughput phage display and deep sequencing. *Mol Biosyst.* 2010 Oct;6(10):1782-90. =3.825

Lam HY, Kim PM, Mok J, Tonikian R, Sidhu SS, Turk BE, Snyder M, Gerstein MB. MOTIPS: automated motif analysis for predicting targets of modular protein domains. *BMC Bioinformatics.* 2010 May 11;11:243. =3.029

Mok J, Kim PM, Lam HYK, Piccirillo S, Zhou X, Jeschke GR, Sheridan DL, Parker SA, Desai V, Jwa M, Cameroni E, Niu H, Good M, Remenyi A, Ma JLN, Sheu YJ, Sassi HE, Sopko R, Chan CSM, De Virgilio C, Hollingsworth NM, Lim WA, Stern DF, Stillman B, Andrews BJ, Gerstein MB, Snyder M, Turk B. Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation motifs. *Sci. Signaling* (2010) Feb 16;3(109):ra12. =6.354

Han SJ, Kim PM. Chaperonin activity modulates codon adaptation. *Mol. Syst. Biol.* (2010);6:342. Epub 2010 Jan 19. =9.667

Costanzo M*, Baryshnikova A*, Bellay J, Kim Y, Spear ED, Sevier SD, Ding H, Koh JLY, Toufighi K, Mostafavi S, Prinz J, St. Onge R, Vandersluis B, Alizadeh S, Bahr S, Brost RL, Chen Y, Cokol M,

Deshpande R, Li Z, Li ZY, Liang W, Marback M, Paw J, San Luis BJ, Shuteriqi E, Hin A, Tong AHY, van Dyk N, Wallace IM, Whitney JA, Weirauch MT, Zhong G, Zhu H, Houry W, Brudno M, Ragibizadeh S, Papp B, Roth FP, Giaever G, Nislow C, Troyanskaya OG, Bussey H, Bader GD, Gingras AC, Morris QD, **Kim PM**, Kaiser CK, Myers CM, Andrews B, Boone C. The Genetic Landscape of the Cell. *Science* (2010) Jan 22;327(5964):425-31. =31.377

Lam HYK*, Mu XJ*, Tanzer A, Stuetz A, Snyder M, **Kim PM**, Korbel JO, Gerstein MB. BreakSeq: Mining a breakpoint library rapidly identifies structural variants in personal genomes and reveals biases in their formation. *Nature Biotech.* (2010) Jan;28(1):47-55. Epub 2009 Dec 27. =31.09

Tonikian R*, Xin X*, Toret CP*, Gfeller D, Landgraf C, Panni S, Paoluzi S, Castagnoli L, Currell B, Seshagiri S, Yu H, Winsor B, Vidal M, Gerstein MB, Bader GD, Volkmer R#, Cesareni G#, Drubin DG#, **Kim PM**#, Sidhu SS#, Boone C#. Bayesian modeling of the yeast SH3 domain interactome predicts spatiotemporal dynamics of endocytosis proteins. *PLoS Biology* 2009 Oct;7(10):e1000218. [#joint corresponding author]. =12.472

Yip KY, **Kim PM**, McDermott D, Gerstein MB. Multi-level Learning: Improving the Prediction of Protein, Domain and Residue Interactions by Allowing Information Flow between Levels. *BMC Bioinformatics* 2009 Aug 5;10:241. =3.029

Alexander RP, **Kim PM**, Emonet T, Gerstein MB. Understanding modularity in molecular networks requires dynamics. *Science Signal.* 2009 Jul 28;2(81):pe44. =6.354

Kim PM*, Lam HY*, Urban AE, Korbel JO, Affourtit J, Grubert F, Chen X, Weissman S, Snyder M, Gerstein MB. Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history. *Genome Res.* 2008 18(12):1865-74. =13.588

Hasin Y, Olender T, Khen M, Gonzaga-Jauregui C, **Kim PM**, Urban AE, Snyder M, Gerstein MB, Lancet D, Korbel JO. High-resolution copy-number variation map reflects human olfactory receptor diversity and evolution. *PLoS Genet.* 2008 Nov;4(11):e1000249. =9.543

Korbel JO*, **Kim PM***, Chen X, Urban AE, Weissman S, Snyder M, Gerstein MB. The current excitement about copy-number variation: how it relates to gene duplications and protein families. *Curr Opin Struct Biol.* 2008 Jun;18(3):366-74. =9.903

Kim PM*, Sboner A*, Xia Y, Gerstein M. The role of disorder in interaction networks: a structural analysis. *Mol Syst Biol.* 2008;4:179. =9.667

Yip KY, Patel P, **Kim PM**, Engelman DM, McDermott D, Gerstein M. An integrated system for studying residue coevolution in proteins. *Bioinformatics*. 2008 Jan 15;24(2):290-2. =4.877

F. Research Funding

Canada Institute for Health Research, Operating Grant, 2012-2017, \$850,000

Ontario Genomics Institute, SPARK Grant, 2012, \$50,000

Genome Canada (subcontract), 2011-2013, total award \$10,000,000

Natural Science and Engineering Research Council, 2010-2016, Discovery Grant, \$122,000

Ontario Research Fund, Global Leadership in Genomics 2 (subcontract), 2010-2014, total award \$2,200,000

Canada Foundation for Innovation, Leadership Opportunity Fund, 2010-2015, \$233,000

Startup Funds, University of Toronto, from 2009, \$475,000

BIOGRAPHICAL SKETCH: Anne-Claude GINGRAS

A. Name and Affiliation

Dr. Anne-Claude Gingras – Senior Investigator, Samuel Lunenfeld Research Institute; Associate Professor, University of Toronto

B. Academic and Training Background

McGill University, Montreal, QC	Ph.D 1994-2001	Biochemistry, Sonenberg lab
Université Laval, Ste-Foy, QC	B.Sc. 1991-1994	Biochemistry

C. Work Experience

2011-	Senior Investigator, Samuel Lunenfeld Research Institute, Toronto, ON
2006-2011	Assistant Professor, Department of Molecular Genetics, University of Toronto
2005-2011	Investigator, Samuel Lunenfeld Research Institute, Toronto, ON
2002-2005	Postdoctoral fellow, Aebersold lab, Inst. for Systems Biology, Seattle, WA

D. Selected Accomplishments and Honours

2012-	Committee member, Canadian Institutes of Health Research (CIHR), Genomics panel
2011	Invitee committee member, Canadian Institutes of Health Research (CIHR), Genomics panel
2011	Co-editor (with Alexey Nesvizhskii) of a special issue of <i>Proteomics</i> on protein-protein interactions
2011	Consultant (category new investigators/early career) for the International review of the Canadian Institutes of Health Research (CIHR)
2011	2011 Canada's Most Powerful Women : Top 100
2010-2011	Full member of the CCSRI (Canadian Cancer Society Research Institute) Panel F (Signaling)
2010	Director of the Research Training Center; Samuel Lunenfeld Research Institute
2010	Co-organizer (with Laurence Pelletier) of the Annual Samuel Lunenfeld Research Institute retreat
2010	Member of the Canadian Research Society Panel A
2010	Member of the Graduate student Recruitment committee, Department of Molecular Genetics, University of Toronto
2010	Director of the Bioinformatics core platform for the National Technology Platform (CFI; PI = Benoit Coulombe)
2009	Editorial Board, Molecular and Cellular Proteomics
2009-2010	Scientific Officer on the CCSRI (Canadian Cancer Society Research Institute) Panel F (Signaling)
2009	Co-organizer of the International Interactome Initiative (I3) Toronto Workshop (with Benoit Coulombe, IRCM and Tony Pawson, SLRI); co-leader of bioinformatics breakout session (with Pascal Braun, Dana Farber, Harvard Medical School).
2008-2011	Member of the Research highlight advisory panel for Nature Reviews Molecular Cell Biology
09/2008	Lea Reichmann Chair in Cancer Proteomics
2008	Scientific Advisory Committee for the University of California in San Francisco (UCSF) mass spectrometry facility
2008-2011	Member of the Canadian Institutes of Health Research (CIHR) Institute of Genetics New Principal Investigator Meeting Organizing Committee
2006-	Co-director (with T. Pawson) of the mass spectrometry platform at the Samuel Lunenfeld Research Institute

Selected invited talks

09/2012	HUPO 2012: 11 th Annual World Congress. Invited plenary speaker.
---------	---

- 08/2012 6th European Summer School: FEBS advanced lecture course: high performance proteomics, Kloster Neustift, Brixen/Bressanone, Southern Tyrol, Italy. Invited lecturer.
- 08/2012 The 13th Conference on Systems Biology, Toronto, ON. Invited speaker and session chair.
- 07/2012 FASEB Summer conference: Tyrosine Kinase Signaling in Cancer, Disease and Development, Snowmass, Colorado. Invited speaker.
- 07/2012 FASEB Summer conference: Protein Phosphatases, Snowmass, Colorado. Invited speaker
- 05/2012 University of Washington, Seattle, Genome Centre. Seminar Series speaker.
- 05/2012 Thermo Fisher mass spectrometry software group, San Jose. Guest speaker.
- 05/2012 AB SCIEX user meeting, ASMS, Vancouver. Invited speaker.
- 05/2012 American Society for mass spectrometry (ASMS) 2012 meeting, Vancouver. Presentation selected from the abstracts.
- 04/2012 RECOMB Satellite Conference on Computational Proteomics 2012, invited speaker.
- 04/2012 2012 meeting on Computational Proteomics, Center for Computational Mass Spectrometry and Center for Algorithmic and Systems Biology, University of California in San Diego, invited speaker.
- 03/2012 US HUPO, 8th Annual Conference, "From Genes to Function". San Francisco, invited speaker.
- 01/2012 8th International Barbados Workshop: from molecular cell biology to clinical proteomics, invited speaker.
- 11/2011 The University of Chicago, Biochemistry and Molecular Biology Seminar Series, invited speaker.
- 09/2011 Protein Modules and Networks in Health and Disease, FEBS Workshop of the Protein Modules Consortium. Seefeld, Austria, invited speaker.
- 08/2011 Tenth International Symposium on Mass Spectrometry in the Health and Life Sciences: Molecular and Cellular Proteomics. San Francisco, California, invited speaker.
- 07/2011 Europhosphatases 2011: Protein phosphatases from molecules to networks; EMBO conference series. Baden, Austria, invited speaker and session chair.
- 07/2011 Department of Pharmaceutical Sciences, UCSF, invited speaker
- 05/2011 Keystone Symposium: Omics meets cell biology, Alpbach, Austria, invited speaker
- 04/2011 HUPO PSI meeting, Heidelberg, invited speaker
- 04/2011 Proteomic Forum 2011, Berlin, invited speaker
- 03/2011 9th Conference on Signaling in Normal and Cancer Cells; Banff, AB, invited speaker and session chair
- 01/2011 Keystone Symposium: The evolution of protein phosphorylation, invited speaker and session chair
- 01/2011 Netherlands Proteomics Center 2011 Progress Meeting, invited speaker
- 08/2010 16th Annual Protein Phosphorylation and Cell Signaling meeting – Salk Institute, La Jolla, CA.
- 06/2010 Gordon Research Conference – Phosphorylation and G-Protein Mediated Signaling Networks, Biddeford, ME.

03/2010	Vanderbilt University, Nashville, TN, invited speaker
10/2009	Case Western Reserve University Symposium: Understanding Protein Complexes, Networks and the Interactome, invited speaker
09/2009	Human Proteome Organization (HUPO) VIII World Congress. Speaker and session chair.
07/2009	ETH Zurich (Switzerland), invited speaker
07/2009	Europhosphatases 2009. Protein phosphatases in development and disease, Egmund ur See, The Netherlands, invited speaker
07/2009	Center for Cancer Systems Biology (CCSB), Dana Farber Institute, Boston, MA. Seminar series speaker
04/2009	Protein interaction workshop, Barbados
11/2008	Pathology Research Seminar Series, University of Michigan, Ann Arbor, MI. invited speaker
11/2008	Angioma Alliance, Pathology of CCM workshop, Washington, DC
01/2008	University of Southern Alabama, invited speaker

E. Peer Reviewed Publications since 2008 (from 79 in career, cited 10697 times; h-index 45)

* indicates co-first authorship; ** indicates corresponding author; Gingras lab members underlined CA = collaborator , PA = principal author, SRA = senior author, IF = impact factor (2010).

1. Nakada, S., Chen, G.I., **Gingras, A.-C.**** and Durocher, D.** (2008) PP4 is a gH2AX phosphatase required for the recovery from the DNA damage checkpoint. *EMBO Reports*, **9**:1019-26. (**co-SRA**). IF=7.822
2. Rong, L., Livingstone, M., Sukarieh, R., Petroulakis, E., **Gingras, A.-C.**, Crosby, K., Smith, B., Polakiewicz, R.D., Pelletier, J., Ferraiuolo, M.A., and Sonenberg, N. (2008) Control of eIF4E cellular localization by eIF4E-binding proteins, 4E-BPs. *RNA* **14**:1318-27. (**CA**). IF=6.051
3. Chen, G.I., Tisayakorn, S., Jorgensen, C., D'Ambrosio, L.M., Goudreault, M. and **Gingras, A.-C.**** (2008) PP4R4/KIAA1622 forms a novel stable cytosolic complex with phosphoprotein phosphatase 4. *J Biol Chem*, **283**:29273-84. (**SRA**). IF=5.328
4. Goudreault, M., D'Ambrosio, L.M., Kean, M.J., Mullin, M., Larsen, B.G., Sanchez, A., Chaudhry, S., Chen, G.I., Sicheri, F., Nesvizhskii, A.I., Aebersold, R., Raught, B., and **Gingras, A.-C.**** (2009) A PP2A phosphatase high-density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. *Mol Cell Proteomics*, **8**:157-71. (**SRA**) IF=8.354
5. Lawo, S., Bashkurov, M., Mullin, M., Gomez Ferreria, M., Kittler, R., Habermann, B., Tagliaferro, A., Poser, I., Hutchins, J., Buchholz, F., Peters, J-M., Hyman, A.A., **Gingras, A.-C.**, and Pelletier, L. (2009) The Augmin complex regulates centrosome and spindle integrity in mammalian cells. *Curr Biol*, **19**:816-26. (**CA**) IF=10.026
6. Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S, Prinz J, St Onge RP, VanderSluis B, Makhnevych T, Vizeacoumar FJ, Alizadeh S, Bahr S, Brost RL, Chen Y, Cokol M, Deshpande R, Li Z, Lin ZY, Liang W, Marback M, Paw J, San Luis BJ, Shuteriqi E, Tong AH, van Dyk N, Wallace IM, Whitney JA, Weirauch MT, Zhong G, Zhu H, Houry WA, Brudno M, Ragibizadeh S, Papp B, Pál C, Roth FP, Giaever G, Nislow C, Troyanskaya OG, Bussey H, Bader GD, **Gingras A.-C.**, Morris QD, Kim PM, Kaiser CA, Myers CL, Andrews BJ, Boone C. (2010) The genetic landscape of a cell. *Science*, **327**:425-21. (**CA**) IF=31.377
7. Bidinosti M., Ran I., Sanchez-Carbente M.R., Martineau Y., **Gingras A.-C.**, Gkogkas C., Raught B., Bramham C.R., Sossin W.S., Costa-Mattioli M., DesGroseillers L., Lacaille J.C., Sonenberg N. (2010) Postnatal deamidation of 4E-BP2 in brain enhances association with raptor and alters kinetics of excitatory synaptic transmission. *Mol Cell*, **37**:797-808. (**CA**) IF=14.194

- 8.** Breitkreutz A., Choi H., Sharon J., Boucher L., Neduva V., Larsen B.G., Lin Z.-Y., Breitkreutz B.-J., Stark C., Liu G., Ahn, J., Dewar-Darch, D., Tang X., Almeida, V., Qin, Z.S., Pawson, T., **Gingras, A.-C.****, Nesvizhskii, A.**, Tyers, M.** (2010) A global protein kinase and phosphatase network. *Science*, **328**:1043-6. (**co-SRA**) IF=31.377
- 9.** Mak, A.B., Ni, Z., Hewel, J., Chen, G.I., Zhong, G., Karamboulas, K., Blakely, K., Smiley, S., Marcon, E., Roudeva, D., Li, J., Olsen, J., Punna, T., Isserlin, R., Chetyrkin, S., **Gingras, A.-C.**, Emili, A., Greenblatt, J. and Moffat, J. (2010) A lentiviral-based functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency. *Mol Cell Proteomics*, **9**:811-23. (**CA**) IF=8.354
- 10.** Choi, H, Kim, S., **Gingras, A.-C.** Nesvizhskii, A.I. (2010). Analysis of protein complexes through model-based biclustering of label-free quantitative AP-MS data. *Mol Syst Biol*, **6**:385. (**CA**) IF=9.667
- 11.** Sydorskyy, Y., Sri Kumar, T., Jeram, S.M., Wheaton, S., Vizeacoumar, F.J., Makhnevych, T., Chong, Y.T., **Gingras, A.-C.** and Raught, B. (2010). A novel mechanism for SUMO system control: regulated Ulp1 nucleolar sequestration. *Mol Cell Biol*, **30**:4452. (**CA**) IF=6.188
- 12.** Nakada, S., Tai, I., Panier, S., Iemura, S.-I., Kumakubo, A., Munro, M., **Gingras, A.-C.**, Natsume, T., Suda, T. and Durocher, D. (2010) Non-canonical inhibition of DNA damage-dependent ubiquitylation by OTUB1. *Nature* 2010 **466**:941-6. (**CA**) IF=36.104
- 13.** Liu, G., Zhang, J.P., Larsen, B., Stark, C., Breitkreutz, A., Lin, Z.-Y., Breitkreutz, B.-J., Ding, Y., Colwill, K., Pascalescu, A., Pawson, T., Wrana, J., Nesvizhskii, A.I., Raught, B., Tyers, M.**, and **Gingras, A.-C.**** (2010) ProHits: an integrated software platform for mass spectrometry-based interaction proteomics. *Nat Biotech*, **28**:1015-7. (**co-SRA**) IF=31.09
- 14.** Baryshnikova A, Costanzo M, Kim Y, Ding H, Koh J, Toufighi K, Youn JY, Ou J, San Luis BJ, Bandyopadhyay S, Hibbs M, Hess D, **Gingras A.-C.**, Bader GD, Troyanskaya OG, Brown GW, Andrews B, Boone C, Myers CL. (2010) Quantitative analysis of fitness and genetic interactions on a genome scale. *Nat Methods*, **7**:1017-24 (**CA**) IF=20.721
- 15.** O'Donnell L, Panier S, Wildenhain J, Tkach JM, Al-Hakim A, Landry MC, Escribano-Diaz C, Szilard RK, Young JT, Munro M, Canny MD, Kolas NK, Zhang W, Harding SM, Ylanko J, Mendez M, Mullin M, Sun T, Habermann B, Datti A, Bristow RG, **Gingras A.-C.**, Tyers M, Brown G and Durocher D. (2010) The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Mol Cell*, **40**:619-31. (**CA**) IF=14.194
- 16.** Choi, H., Larsen, B., Lin., Z.-Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z.S., Tyers, M., **Gingras, A.-C.**** and Nesvizhskii, A.I.** (2011) SAINT: probabilistic scoring of affinity purification - mass spectrometry data. *Nature Methods*, **8**:70-3. (**co-SRA**) IF=20.721
- 17.** Templeton, G., Nimixk, M., Morrice, N.A., Campbell, D.G., Goudreault, M., **Gingras, A.-C.**, Takemiya, A., Shimazaki, K.I., Moorhead, G.B. (2011) Identification and characterization of Atl-2, an Arabidopsis homolog of an ancient protein phosphatase (PP1) regulatory subunit. *Biochem J*, **435**:73-83. (**CA**) IF=5.016
- 18.** Skarra, D.V., Goudreault, M., Choi, H., Mullin, M., Nesvizhskii, A., **Gingras, A.-C.****, and Honkanen, R.** (2011) Label-free quantitative proteomics and SAINT analysis enable interactome mapping for the human Ser/Thr protein phosphatase 5. *Proteomics*, **11**:1508-16. (**co-SRA**) IF=4.815
- 19.** Li, Z., Vizeacoumar, F., Bahr, S., Li, J., Warringer, J., Vizeacoumar, F., VanderSluis, B., Bellay, J., DeVit, M., Fleming, J., Stephens, A., Haase, J., Lin, Z.-Y., Baryshnikova, A., Min, R., Lu, H., Yan, Z., Jin, K., Barker, S., Datti, A., Giaever, G., Nislow, C., Bulawa, C., Costanzo, M., Myers, C., **Gingras, A.-C.**, Zhang, Z., Blomberg, A., Bloom, K., Andrews, B. and Boone, C. (2011) Systematic exploration of essential gene function with temperature-sensitive mutants. *Nature Biotech*, **29**:361-7. (**CA**) IF=31.09
- 20.** Dunham, W., Larsen, B., Tate, S., Gonzalez Badillo, B., Tehami, Y., Kislinger, T., Goudreault M and **Gingras, A.-C.**** (2011) A cost-benefit analysis of multidimensional fractionation of affinity purification – mass spectrometry samples. *Proteomics*, **11**:2603-12. (**SRA**) IF=4.815
- 21.** Kean, M.J., Ceccarelli, D., Goudreault.. M., Tate, S., Larsen, B., Sanches, M., Gibson, L.C., Derry., W.B., Scott, I.C., Pelletier, L., Baillie, G.S., Sicheri, F., and **Gingras, A.-C.**** (2011) Structure-function analysis of core STRIPAK proteins: a signaling complex implicated in Golgi polarization. *J Biol Chem*, **15**:25065-75. (**SRA**) IF=5.328

- 22.** Ceccarelli, D.F., Laister, R.C., Mulligan V.K., Kean, M.J., Goudreault, M., Scott, I.C., Derry, W.B., Chakrabart A., **Gingras, A.-C.**** and Sicheri, F.** (2011) CCM3/PDCD10 Heterodimerizes with germinal center kinase III (GCKIII) proteins using a mechanism analogous to CCM3 homodimerization. *J Biol Chem*, **15**:25056-64. (**co-SRA**) IF=5.328
- 23.** Olhovsky, M., Williton, K., Dai, A.Y., Pascalescu, A., Lee, J.P., Goudreault, M., Wells, C.D., Park, J.G., **Gingras, A.-C.**, Linding, R., Pawson, T., and Colwill, K. (2011) OpenFreezer: A reagent information management software system. *Nature Methods*, **28**:612-3. (**CA**) IF=20.721
- 24.** Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., **Gingras, A-C.**, Arrowsmith, CH., and Knapp, S. (2012) Histone recognition and large structural analysis of the human bromodomain family. *Cell*, **149**(1):214-31. (**CA**) IF=32.406
- 25.** Al-hakim, A.K., Bashkurov, M., **Gingras, A.-C.**, Durocher, D and Pelletier, L. (2012) Interaction proteomics identify NEURL4 and the HECT E3 ligase HERC2 as novel modulators of centrosome architecture. *Mol Cell Proteomics, Epub PMID:22261722* (**CA**) IF=8.354
- 26.** Knight, J.R., Tian, R., Lee, R.E.C., Wang, F., Beauvais, A., Zou, H., Megeney, L.A., Gingras, A.-C., Pawson, T., Figeys, D. and Kothary, R. (2012) A novel whole-cell lysate kinase assay identifies substrates of the p38 MAPK in differentiating myoblasts. *Skeletal Muscle*, **2**(1):5. (**CA**)
- 27.** Jovic, M., Kean, M.J., Szentpetery, Z., Polevoy, G., **Gingras, A.-C.**, Brill, J.A. and Balla, T. (2012) Two PI 4-kinases control lysosomal delivery of the Gaucher disease enzyme, β -glucocerebrosidase. *Mol Biol Cell*, **23**(8):1533-45. (**CA**) IF=5.861
- 28.** Gomez-Ferreria, M.A., Bashkurov, M., Helbig, A., Larsen, B., Pawson, T., **Gingras, A.-C.** and Pelletier, L. (2012) NEDD1 phosphorylation is modulated by CEP192 and regulates its binding to γ -tubulin. *J Cell Science*, May 17 (**CA**) IF=6.29
- 29.** van Zuijlen, W., Doyon, P., Clément, J.-F., Khan, K.A., D'Ambrosio, L.M., Dô, F., St-Amant-Verret, M., Wissanji, T., Emery, G., **Gingras, A.-C.**, Meloche, S., and Servant, M.J. (2012) Proteomic profiling of the TRAF3 interactome network reveals a new role for the ER-to-Golgi transport compartments in innate immunity. *PLOS Pathogens*, **8**(7):e1002747. (**CA**) IF=9.079
- 30.** Belozerov, V.E., Lin, Z.-Y., **Gingras, A.-C.**, McDermott, J., and Siu, K.W.M. (2012) High-Resolution Protein Interaction Map of the *Drosophila melanogaster* p38 Mitogen-Activated Protein Kinases Reveals Limited Functional Redundancy. *Mol Cell Biol*, **32**(18):3695-706. (**CA**) IF=6.188
- 31.** Costa, B., Kean, M.J., Ast, V., Knight, J.R., Mett, A., Levy, Z., Ceccarelli, D., Gonzales Badillo, B., Eils, R., König, R., **Gingras, A.-C.** and Fainzilber, M. (2012) STK25 protein mediates TrkA and CCM2 protein-dependent death in pediatric tumor cells of neural origin. *J Biol Chem*, **24**:287(35):29285-9. (**CA**) IF=5.328
- 32.** Morita, M., Ler, L.W., Fabian, M.R., Siddiqui, N., Mullin, M., Henderson, V., Alain, T., Fonseca, B.D., Karashchuk, G., Bennett, C., Kabuta, T., Higashi, S., Larsson, O., Topisirovic, I., Smith, R.J., **Gingras, A.-C.** and Sonenberg, N. (2012) A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. *Mol Cell Biol*, **32**(17):3585-93. (**CA**) IF=6.188
- 33.** Ryan, O., Shapiro, R.S., Kurat, C.F., Mayhew, D., Baryshnikova, A., Chin, B., Lin, Z.-Y., Cox, M.J., Vizeacoumar, F., Cheung, D., Tsui, K., Istel, F., Schwarzmuller, T., Reynold, T.B., Kuchler, K., Gifford, D.K., Giaever, G., Nislow, C., Costanzo, M., **Gingras, A.-C.**, Mitra, R.D., Andrews, B., Fink, G.R., Cowen, L.E., and Boone, C. (2012) Global Gene Deletion Analysis Exploring Yeast Filamentous Growth. *Science*, **337**(6100):1353-6. (**CA**) IF=31.377
- 34.** Gomez-Ferreria, M.A., Bashkurov, M., Mullin, M., **Gingras, A.-C.** and Pelletier, L. (2012) CEP192 interacts physically and functionally with the K63-deubiquitinase CYLD to promote mitotic spindle assembly. *Cell Cycle*, **16**:11(19). (**CA**) IF=4.999

F. Review or methods articles since 2008 (from 24 in career; most are peer-reviewed)

- 1.** **Gingras, A.-C.**** (2009) Journal Club: 35 later, mRNA caps still matter. *Nature Reviews Mol Cell Biol*, **10**:734. (**SRA**) IF=38.65

2. **Gingras, A.-C.**** (2011) Protein Phosphatases, from Molecules to Networks. *EMBO Reports*, **12**:1211-13. (**PA, SRA**) IF=7.822
3. Lambert, J.-P., Pawson, T. and **Gingras, A.-C.**** (2012) Mapping physical interactions within chromatin by proteomic approaches. *Proteomics*, **12**(10):1609-22. (**SRA**) IF=4.815
4. Dunham, W., Mullin, M and **Gingras, A.-C.**** (2012) Affinity-Purification coupled to Mass Spectrometry: Basic Principles and Strategies. *Proteomics*, **12**(10):1576-90. (**SRA**) IF=4.815
5. St-Denis, N., and **Gingras, A.-C.**** (2012) Mass spectrometric tools for systematic analysis of protein phosphorylation. *Progress in molecular biology and translational sciences: Protein Phosphorylation in Health and Disease*, Elsevier, edited by S. Shenolikar, **106**:3-32. (**SRA**) IF=3.92
6. Braun, P. and **Gingras, A.-C.** (2012) Protein interactions in the 20th century: from egg white to complex networks. *Proteomics*, **12**(10):1478-98. (**CA**) F=4.815
7. **Gingras, A.-C.**** and Raught, B. (2012) Beyond hairballs: the use of quantitative mass spectrometry data to understand protein-protein interactions. *FEBS Letters*, **586**(17):2723-31. (**PA, SRA**) IF=3.601
8. Liu, G., Zhang, J., Choi, H., Lambert, J.-P., Srikumar, T., Larsen, B., Nesvizhskii, A.I., Raught, B., Tyers, M., and **Gingras, A.-C.**** (2012) Using ProHits to store, annotate and analyze affinity purification - mass spectrometry (AP-MS) data. *Current Protocols in Bioinformatics*, **8**(8):16. (**SRA**)
9. Choi, H., Liu, G., Tyers, M., **Gingras, A.-C.**** and Nesvizhskii, A.I.** (2012) Analyzing protein-protein interactions from affinity purification-mass spectrometry data with SAINT. *Current Protocols in Bioinformatics*, **8**(8):15. (**co-SRA**)
10. Kean, M.J., Couzens, A.L. and **Gingras, A.-C.**** (2012) Mass spectrometry approaches to study mammalian kinase and phosphatase associated proteins. *Methods*, **57**(4):400-8. (**SRA**)
11. Tate, S., Larsen, B., Bonner, R and **Gingras, A.-C.** (2012) Quantitative proteomics trends for protein-protein interactions. *Accepted with minor modifications in The Journal of Proteomics*. (**CA**) IF=5.074

G. Active Research Funding

- Source: Canadian Institutes for Health Research Operating Grant
 Title: Global approaches to unravel PP2A function
 PI: Gingras, Anne-Claude
 Dates: 2010-2015
 Amount: 192077\$/year
- Source: Canadian Cancer Society Research Institute (CCSRI)
 Title: Structure, function and regulation of PP4cs
 PI: Gingras, Anne-Claude
 Dates: 2009-2014
 Amount: 137721\$/year
- Source: Ontario Research Fund – Global Leadership Round in Genomics
 Title: Understanding the assembly and function of dynamic signalling networks in complex diseases
 PI: Pawson, A.J (Gingras is co-PI)
 Dates: 2010-2015
 Amount: 179634\$/year to AC Gingras
- Source: National Institute of Health (NIH) – RO1
 Title: Computational tools for mass spectrometry-based Interactome
 PI: Nesvizhskii, Alexey (Gingras is co-PI)
 Dates: 2010-2015
 Amount: 50000\$/year to AC Gingras
- Source: CIHR Institute of Genetics – Bridge Funding

Title: Molecular Mechanisms of Cerebral Cavernous Malformations
PI: Gingras, A.-C. (Brent Derry is co-PI)
Dates: April 2012-March 2013
Amount: 60000\$ to AC Gingras (total 100000\$)

Source: Canadian Institutes for Health Research Operating Grant
Title: Molecular Mechanisms of Cerebral Cavernous Malformations
PI: Gingras, A.-C. (Brent Derry is co-PI)
Dates: August 2012 – July 2017
Amount: 105000\$/year to AC Gingras (total 190026\$/year)

Source: Canadian Institutes for Health Research Operating Grant
Title: A systems approach towards the therapeutic modulation of the acetylome
PI: Gingras, A.-C. (Panagis Filippakopoulos and Tony Pawson are co-PIs)
Dates: Oct 2012-Sept 2016
Amount: 120000\$/year to AC Gingras (total 202115\$/year)


Alexey I. Nesvizhskii, Ph.D.

Associate Professor,
 Department of Pathology
 Department of Computational Medicine & Bioinformatics
 University of Michigan Medical School
 Email: nesvi@umich.edu
 Phone: 734-764-3616
 Fax: 734- 936-7361

September 23, 2012

To: Anne-Claude Gingras
 Senior Investigator, Samuel Lunenfeld Research Institute at Mount Sinai Hospital
 Associate Professor, Department of Molecular Genetics, University of Toronto

Dear Anne-Claude,

I am writing this letter in full support of the CCSRI grant entitled “Rapid mapping of cancer interactomes” for which you will be profiling interactome changes imparted by driver mutations in cancers, as well as the effect of therapeutic drugs on the modulation of protein-protein interactions. I believe that your project is not only innovative in the context of finding out the mechanisms of tumor formation, but that the methods that you will be developing within the context of this project will also open up new avenues of investigation for researchers working directly with patient samples. We have an excellent track-record at collaborating with each other to solve key problems in proteomics bioinformatics, particularly with regards to the development of tools for interaction proteomics, and your proposed project will offer excellent opportunities to extend our collaborations.

In particular, the data independent acquisition strategy (SWATH) that you are proposing is very innovative in that it should enable robust and accurate label-free quantification in a scalable manner. Since the approach is so new (with the first manuscript describing its use in the same set-up as you published in 2012), your leadership on this project, and the multidisciplinary collaborations you have established with many groups, including us and the mass spectrometry vendor AB SCIEX, places you in an ideal position to introduce new paradigms in the way researchers analyze protein samples by mass spectrometry. While there has been some debate regarding the application of data independent acquisition to very complex samples, the affinity-purified samples that you are generating within the context of your grant are perfectly appropriate to quantification by SWATH mass spectrometry. I am very excited by the preliminary data (proof-of-principle) on the CDK4 point mutants, the GRK6 splice variants and the interactions modulated by the HSP90 inhibitor which you are including in the manuscript by Lambert et al. to be submitted shortly. In particular, I was impressed that by employing methodology designed for quantification of SRM (Selected Reaction Monitoring) data, you were already able to extract excellent quantitative information in a timely manner. These results can only be improved upon, and I am looking forward to work with you on this within the context of your CCSRI project.

My group has already begun further analyzing the data generated by your group on these samples, and that we are working on approaches to further improve the data analysis pipeline for data independent acquisition. In particular, we are currently (with my new student Chih-Chiang Tsou) investigating whether we can employ the SWATH data for simultaneous identification and quantification of the peptides/proteins present in your samples. This would be a significant change to the current pipeline in which you are creating a separate spectral library using standard data-dependent acquisition which you are using for extracting quantitative information from the SWATH (data independent acquisition) runs. In general, this approach should increase your throughput significantly by decreasing the instrument acquisition time, which is a limiting factor at the moment. We are also working on further characterizing with you important aspects of SWATH analysis, including determining

the limit of detection, influence of sample complexity on both identification and quantification, and the dynamic range for quantification. I need to stress that these issues do not immediately impinge on the type of affinity purified samples which you are proposing in the current application, but will have an impact on future studies by your team and others, in which samples of different composition(s) will need to be analyzed by data independent mass spectrometry.

In addition, the committee should know that I have applied for a supplement to our joint NIH grant to generate statistical scoring tools for interaction mapping using SWATH data. In the parent NIH proposal, we co-developed the Significance Analysis of INTeractome (SAINT) software tool which is critical in ensuring detection and reporting of only high quality protein interaction data. While we developed SAINT initially for cruder quantification approaches, adapting it to the analysis of SWATH data will be enabled in the near future. As usual, I will be happy to assist your group in implementing all the tools we are co-developing within your excellent data management system “ProHits”.

My group is also developing additional tools for the analysis of proteomics data, which I am making fully available to you. For the interactomes produced directly from cancer cells (which you have already begun investigating), a possible difficulty in data analysis comes from the fact that there may be a disconnect between the consensus sequence(s) present in the protein databases (the first step in identifying a peptide relies on genomics data, namely in the generation of conceptually translated databases that are used to general theoretical spectra for matching purposes). I am therefore working on a number of approaches that will utilize data directly generated from DNA and RNA sequencing projects to create specific protein databases for different tumors / cell types, and I believe that this should greatly facilitate analysis of tumor samples by proteomics. These tools will be built within the time-frame of your project and should enable more efficient analysis, especially as you move beyond the pilot screens in model cell lines towards patient-derived materials. Together with a number of other specialized bioinformatics tools that we are developing (for example to resolve issues with quantification of peptides shared in sequence identity across many proteins or for better assessing the exact site of a post-translational modification), these new resources should enable you to gain the maximum of high quality information from your datasets.

I am looking forward to our continued collaboration. Good luck with the application,



Alexey I. Nesvizhskii, Ph.D.
Associate Professor,
University of Michigan, Ann Arbor



DANA-FARBER
CANCER INSTITUTE



Marc Vidal, PhD

Director
Center for Cancer Systems Biology (CCSB)
Professor of Genetics
Harvard Medical School
marc_vidal@dfci.harvard.edu



David E Hill, PhD

Associate Director
Center for Cancer Systems Biology (CCSB)
Principal Research Scientist
Department of Cancer Biology
david_hill@dfci.harvard.edu

Dana-Farber Cancer Institute
450 Brookline Ave
Boston, MA 02215
<http://ccsb.dfci.harvard.edu>
617.632.5114 tel

Boston, September 27, 2012

Anne-Claude Gingras

Senior Investigator, Samuel Lunenfeld Research Institute at Mount Sinai Hospital
Associate Professor, Department of Molecular Genetics, University of Toronto

Dear Anne-Claude,

This letter is to offer our enthusiastic support for your application to the Canadian Cancer Society Research Institute entitled "Rapid mapping of cancer interactomes", in which you are proposing to use new mass spectrometry tools to quantify changes in protein interactions due to cancer-associated mutations.

As you are very well aware, our Center for Cancer Systems Biology (CCSB) has a long-standing interest in the development of methods and reagents to understand the phenotypical consequences of genetic changes. In particular we, like you, recognize the growing disconnect between the DNA/RNA sequencing efforts (which may lead to >100,000 genome sequences by the end of this decade), and the functional consequences of all the mutations that are being discovered. We argue that this disconnect truly creates a bottleneck in the understanding of cancer initiation, and limits the ability to treat it.

We realized many years ago that a common consequence of many disease-associated mutations was to specifically perturb particular protein-protein interactions,. We have termed this category of sequence variants "edgetic mutants", based on the systems biology description of an interaction as an "edge" that connects two "nodes" (i.e., the proteins involved in the interaction). Since the publication of our manuscripts describing the concept of "edgetic" mutations in 2009, we have created the largest collection of disease-associated mutations in the world, and we have assembled a multi-disciplinary team that will continue collaborating in the study of genotype to phenotype consequences of mutations. We have recently applied for NIH funding for a

Center of Excellence in Genomic Science (CEGS) "Edgotyping Initiative" grant, of which you are a key collaborator.

In particular, we are impressed by the complementarity of the information that can be gained by mapping binary interactions (which is what our group primarily focuses on) and by affinity purification coupled to mass spectrometry as you are doing. We are excited by your newly introduced data-independent acquisition quantitative proteomics mapping, as we believe that this finally offers a scalable solution to AP-MS analysis that is compatible with systematic efforts. This is much needed for tackling interactome projects, and your initial proof-of-principle manuscript (which we are proud to be part of) indicates the great promise of this approach for edgotyping studies.

Your current proposal, which nicely builds on the technology you have implemented, will be facilitated by your continued access to the collection of reagents that our group has been generating for our ongoing edgotyping efforts. In particular, we are creating a collection of allelic variants for disease-associated genes in a Gateway Entry vector that we will make available to you for your analytical pipeline (We know that you are using Gateway-mediated cloning for your own projects). And we are looking very much forward to the results that you will obtain.

We are also excited by your plan to provide early access to your AP-SWATH data to us and your other collaborators, as this will greatly accelerate the pace of discovery and help the entire team prioritizing follow-up studies. Similarly, you and our other CEGS Edgotyping Initiative collaborators will have access to our unpublished data.

We are looking forward in having you and your research group as key collaborators in our Edgotyping Initiative. We wish you all the best with your grant application, and are looking forward to a productive and continued collaboration.

Sincerely,


Marc Avidal, PhD
Professor of Genetics


David E Hill, PhD
Principal Research Scientist



September-28-12

Anne-Claude Gingras

Senior Investigator, Samuel Lunenfeld Research Institute
Associate Professor, Dept of Molecular Genetics, U of Toronto

Dear Anne-Claude,

I am excited by your Innovation grant to the Canadian Cancer Society in which you are proposing to use a new mass spectrometric approach to systematically investigate the consequences of point mutations on a protein's ability to interact with its partners. In particular, I was very impressed by your preliminary results that indicate that the approach is quantitative and scalable, which is critical to meet the need for functional characterization of all the cancer-associated mutations that are being discovered thanks to the efforts of the genomics community.

I was particularly thrilled to see your results with the effects of the HSP90 inhibitors on the dissociation of the wild type CDK4 from HSP90 and its kinase-specific adaptor CDC37, and the relative resistance of the mutants of CDK4 from inhibitor-induced dissociation. The results you showed me in which you identified a target of a natural product with cancer therapeutic potential in one of your other datasets and then used SWATH to demonstrate that this product was in fact disrupting specific protein-protein interactions also demonstrate the importance of non-assumptive approaches to rapidly increase the knowledge of therapeutic avenues.

As you know, I have a long-standing interest in chemical biology and high-throughput screening. In particular, I oversee with our colleague Jeff Wrana the SMART facility at the Lunenfeld, which is managed by Alessandro Datti assisted by a staff of three members.

The Lunenfeld SMART facility houses several chemical libraries, including the TOCRIS Bioscience library consisting of ~1400 compounds affecting specific cell signaling pathways and/or cellular processes, other collections such as the Prestwick chemical library, the SpectrumCollection, the OICR kinase inhibitors collection (n=400), and the NIH Clinical Collection (BioFocus, n=450 marketed drug, and others in phase I-III clinical trials). We also have the *Library of Pharmacologically Active Compounds* (LOPAC, 1280 compounds) and Biomol chemical known bioactive compounds, an FDA approved drug library.

Cherry picking and screens are available at cost-back to the academic use, and I will assist you in selecting compounds based on your interaction data.

I am looking very much forward to helping you develop this platform that has the potential to be transformative for cancer biology researchers. Good luck with your CCSRI!

Best Regards,

A handwritten signature in black ink that reads "James W. Dennis".

James W. Dennis PhD
Canada Research Chair
Senior Investigator & Professor
Department of Molecular Genetics
Samuel Lunenfeld Research Institute
Mount Sinai Hospital
600 University Ave. R988



71 Four Valley Drive
Concord, Ontario L4K 4V8
Canada

Section / Title: Applicant info / gingras_anneclaude@bells.ca

October 1, 2012

Anne-Claude Gingras, Ph.D.,
Senior Investigator, Samuel Lunenfeld Research Institute at Mount Sinai Hospital
Associate Professor, Department of Molecular Genetics, University of Toronto

Dear Anne-Claude,

This letter is to offer our enthusiastic support for your CCSRI application on the definition of the interactome changes imparted by cancer-associated mutations using the SWATH data independent acquisition technology. The initial data you have obtained on the mutations of the cyclin dependent kinase CDK4, the splice variants of the GPCR-coupled kinase GRK6 and therapeutic inhibition of the HSP90 chaperone are excellent and exciting examples of the promise of the AP-SWATH approach for analyzing such complex interactions.

The committee should know that we have had a long-standing collaboration with you and your colleagues at the Lunenfeld. We have developed improved approaches for Selected Reaction Monitoring on qTRAP instruments, which have resulted in two publications and a submitted manuscript to date. We have also been collaborating with your group on the testing of our newest generation mass spectrometer, the 5600 TripleTOF, for proteomics datasets; and I was pleased that the first publication worldwide on this instrument came from our concerted efforts (Dunham et al., *Proteomics*, 2011). Since then, we have been actively working with your group to test different modes of quantification on this instrument. These include intensity-based measurements of the precursor ions (Kean et al., *J Biol Chem*, 2011), a modified SRM-like quantification that we call MRM^{HR} (Zheng et al., *under review at Science*), and the SWATH approach that you are proposing to use extensively here (Lambert et al., *in prep*). The data we have obtained together has been very well received at numerous scientific conferences where members of our groups have presented.

Importantly, for several of the recent co-publications between the Lunenfeld and our team, we applied our in-house informatics expertise to build new analytical tools for analyzing MS/MS based quantification data. This enabled the quick analysis of the new type of quantification data using the data dependent acquisition SWATH methodology.

We are fully committed to continuing our joint efforts to facilitate seamless data analysis. You will also, as always, have privileged access to new techniques and tools developed within our research group.

I am looking forward to our continued collaboration, and wish you all the best with your CCSRI application.

Sincerely,

Chris Lock, Ph.D.,
Director, Mass Spectrometry Research
AB Sciex



600 University Avenue, Rm 1084
Toronto, ON, Canada M5G 1X5
www.mshri.on.ca

Professor Anthony J. Pawson, CH, FRS, OC
Distinguished Scientist
Tel: 416-586-8262
Fax: 416-586-8869
Email: pawson@Lunenfeld.ca
website: <http://pawsonlab.mshri.on.ca>

September 28, 2012

Anne-Claude Gingras, Ph.D.
Senior Investigator
Samuel Lunenfeld Research Institute
Mount Sinai Hospital
Associate Professor
Department of Molecular Genetics, University of Toronto

Dear Anne-Claude:

This letter is in full support of your proposal to the Canadian Cancer Society which builds on some recent and exciting data which our two groups have been generating by combining affinity purification and quantitative mass spectrometry.

The committee should know that we have had a productive collaboration since you joined the Lunenfeld where we co-direct the mass spectrometry core, collaborate on several grants, have seven collaborative publications (and several more submitted or in preparation), and co-supervise two postdoctoral fellows, James Knight and Jean-Philippe Lambert. In particular, Jean-Philippe has led the exciting development of the SWATH approach as an unbiased quantitative tool permitting systematic mapping of interactome changes imparted by cancer-associated mutations which forms the basis for your current application.

My group has realized a long time ago that the standard techniques that used mass spectrometry for mapping protein-protein interactions were insufficiently robust to analyze quantitative changes associated with perturbations. In part, this is due to the stochasticity of standard mass spectrometry acquisition techniques, which introduces irreproducibility in measurements, especially when the proteins quantified are present in low abundance or the changes are modest. Though isotopic labeling partially overcame these issues when comparing samples within a single experiment, the problem still arises when comparisons are needed across a larger number of samples.

For this and other reasons, my group has worked on the development of alternative approaches for protein quantification by mass spectrometry. We in fact introduced

recently the concept of coupling affinity purification to a technique known as SRM (Selected Reaction Monitoring) that does not suffer from this stochasticity problem as desired species are continuously monitored. We have used this AP-SRM approach for quantifying changes in the interactions established by signaling scaffolds in the receptor tyrosine kinase (RTK) signaling pathways following treatment with growth factors or inhibitors (Bisson et al., *Nature Biotechnology*, 2011; Zheng et al., under review at *Nature*). We showed in these manuscripts that the quantification is robust and accurate, enabling us in the case of the Zheng manuscript (of which you are an author) to perform time-course measurements revealing the dynamic association of ~30 proteins with the ShcA scaffold.

The AP-SRM approach is robust and accurate, and we are in fact developing a large number of assays based on this technique which we are hoping will improve the assessment of the response of patients to therapeutic agents, but it is not ideally-suited to the analysis of multiple independent interactomes as you are proposing to do here. In fact, there are two points that make AP-SRM less than ideal for your project: 1) building the SRM approach itself is time-consuming, and requires careful optimization of the transitions (the pair of parent ion / product ion), which takes weeks or months per target AP sample; 2) the approach can only monitor those proteins which have been preselected; in your case, this would hamper detection of “gains” in the interactome imparted by mutations.

The AP-SWATH approach, which you have introduced as an alternative, is in fact ideally-suited for systematic interactome studies. It does exhibit many of the same characteristics as SRM in terms of measurement accuracy (as in SRM, quantification occurs at the level of the MS/MS spectrum), and in fact, many of the tools developed for quantification of SRM data can be relatively easily adapted to SWATH analysis, which is what you have done. However, it has two main advantages over SRM for your project: 1) there is no optimization needed at the level of individual biological sample, which means that all the baits can be analyzed using the same parameters; 2) the entire content of the sample is fragmented (MS/MS spectra), enabling in depth characterization of the sample, and analysis of both interaction gain and loss. While I realize that it is still early days for SWATH, the collective data our groups have obtained at the Lunenfeld, alongside optimization in the lab of our collaborator Prof. Aebersold and by our industrial partner AB Sciex, fully comfort me in believing that the AP-SWATH combination is ideally-suited for the project you are proposing here. The collaborations we are both maintaining with other groups actively developing data independent acquisition methods (and in particular your close interaction with the leading groups involved in computational proteomics) will ensure timely delivery on your proposed aims.

Lastly, I am very excited by the perspective of uncovering new biology, and will be happy to continue collaborating with you on follow-up experiments, especially for baits that are closely related to the research interests of my group in signal transduction.

I therefore wish you all the best with your application, and am looking forward to continue our collaborations.

Sincerely,

A handwritten signature in blue ink that reads "Tony Pawson". The signature is fluid and cursive, with "Tony" on top and "Pawson" below it, both starting with a capital letter.

Professor Anthony J. Pawson, CH, FRS, OC

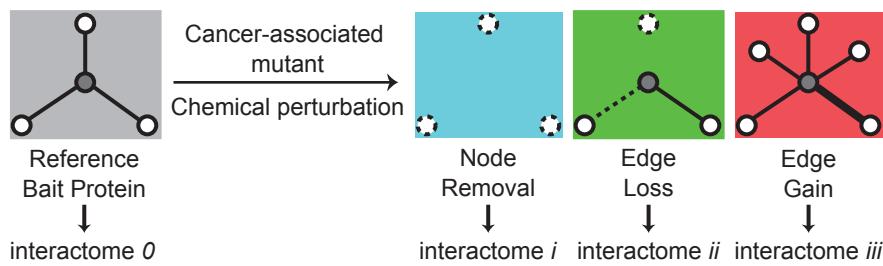


Figure 1. Potential consequences of mutation(s) on interactome changes. The wild type protein (left; “bait” protein) interacts with three proteins. If the bait is completely removed (“node” removal), all interactions are simultaneously lost, resulting in modified interactome *i*. By contrast, some point mutations may alter only some of the interactions, generating different interactomes with different transforming potential.

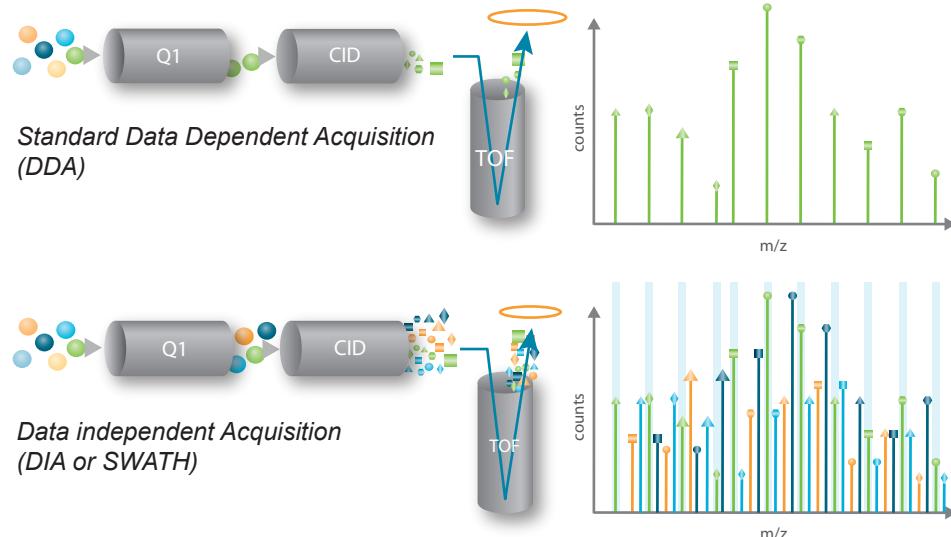


Figure 2. Differences between DDA and DIA mass spectrometry methods. (top) In classical DDA, co-eluting peptides are first analyzed prior to fragmentation in a survey scan (MS1). The *n* most abundant species are then sequentially isolated in the mass spectrometer, fragmented in the CID chamber and the fragment masses are measured in the TOF chamber. A typical MS/MS spectrum is shown on the right: these spectra are used for peptide identification. (bottom) In DIA (here SWATH), there is no decision to select ions for fragmentation based on a precursor scan. Instead, the Q1 filters ions by relatively large mass windows: co-filtered peptides are co-fragmented in the CID chamber and co-analyzed in the TOF chamber. This results in a mixed spectrum, shown on the right, from which identification and quantification information can be retrieved using bioinformatics tools. The process is rapid such that the entire mass range can be analyzed in ~3 seconds.

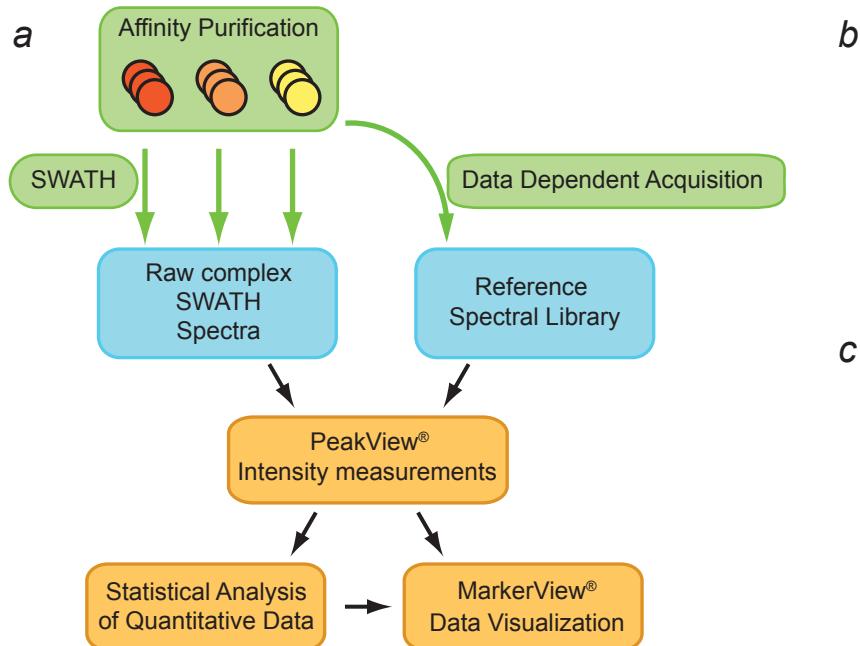
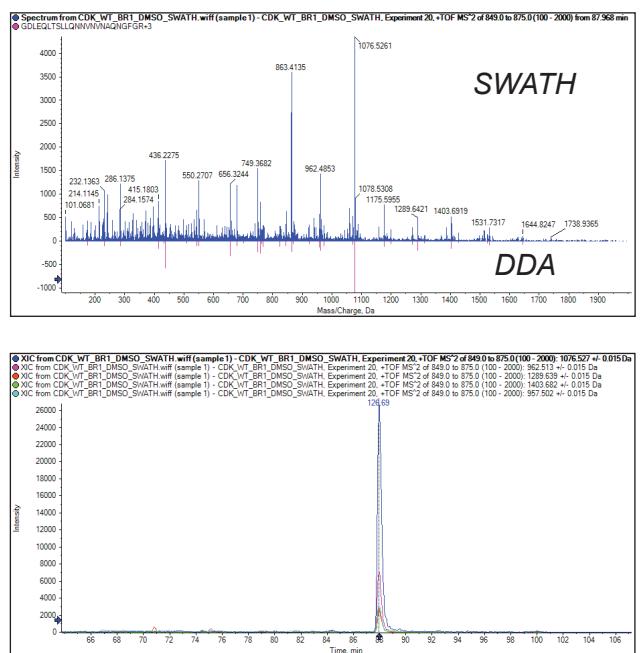


Figure 3. Our current analysis pipeline that couples affinity purification and SWATH (AP-SWATH). (a) General procedures. The wild type and cancer mutant proteins are purified on a FLAG magnetic bead resin in parallel (biological replicates are generated). Half of the samples is analyzed by standard DDA to generate a reference spectral library. The other half is analyzed by SWATH (this will be used for abundance measurements). Software tools from our collaborator AB Sciex are used to (b) match the SWATH spectra (blue) to the Reference spectrum acquired by DDA (pink) and to (c) extract the quantitative information based on the signal intensities of extracted MS/MS fragment ions across the chromatographic window. Subsequently to this data processing, statistical tools are used that help extract confidently regulated interactions.



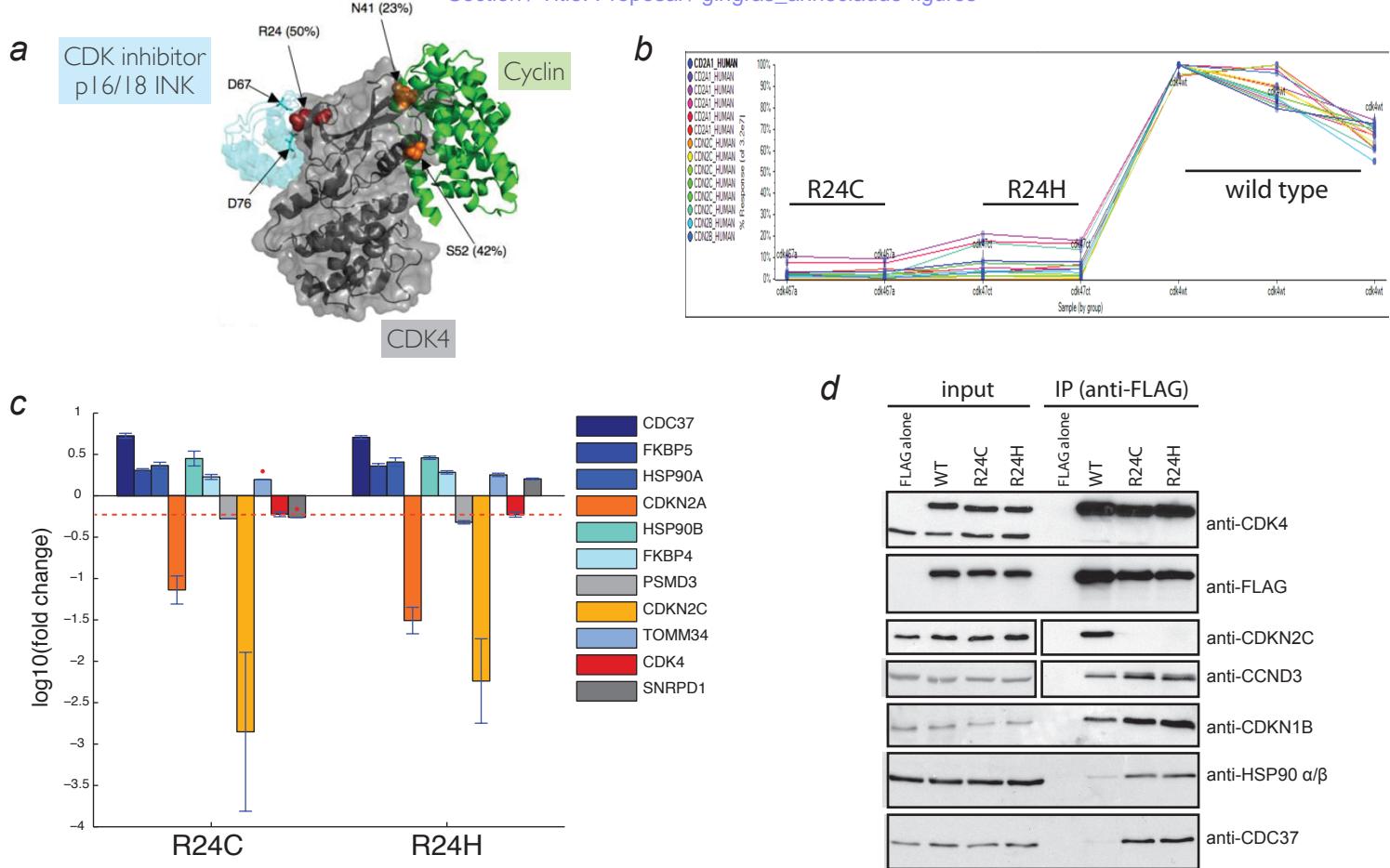


Figure 4. AP-SWATH identifies regulated interactions for melanoma-associated CDK4 mutants. (a) the selected mutants (R24C and R24H) affect the interaction with the INK family of CDK inhibitors. (b) non-normalized intensities for all peptides derived from INK proteins across two biological replicates of each mutant and three replicates of the wild-type: note that the mutants have a much decreased interaction with the INK proteins. (c) Statistical analysis of the entire dataset (expressed in fold change for the mutant over the wt) reveals statistically changed interactors: as expected, the INK proteins are decreased in association with the mutants (orange). By contrast, HSP90 proteins - and in particular the kinase-specific adaptor CDC37 (blue colors) bind more strongly to the mutants. (d) validation of regulated interactions by IP-western.

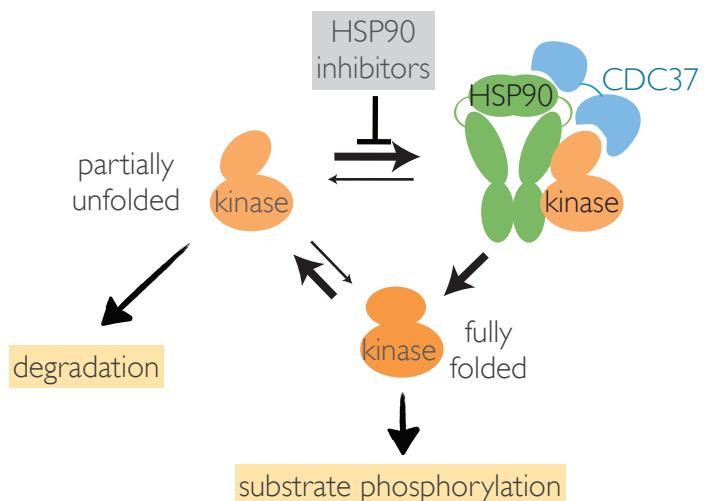


Figure 5. Therapeutic rationale for the use of HSP90 inhibitors for tumors driven by oncogenic kinases: oncogenic kinases display a higher requirement for folding by CDC37/HSP90. Disrupting the recruitment of the kinases to the chaperone ultimately results in their degradation. At odds with the literature on other kinases, the R24 CDK4 mutants dissociated at a lesser extent than the wt CDK4 from CDC37/HSP90 after treatment, as detected by AP-SWATH (not shown).

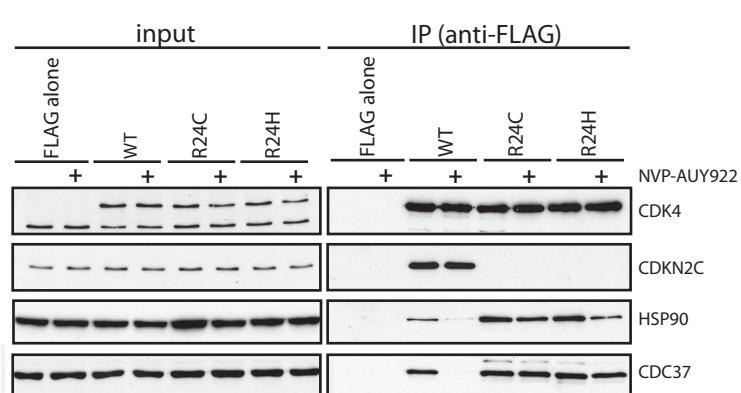


Figure 6. Confirmation by IP-western of the AP-SWATH finding that the R24C and R24H mutants are more resistant than the wt CDK4 to dissociation from CDC37/HSP90 due to a short time treatment with a clinically relevant HSP90 inhibitor. The same phenomenon is observed at later time points, and with higher doses of inhibitor.

Gingras, Anne-Claude

Justification - supplies

EXPENDABLES

We are only requesting expendable for the wet lab experiments which will be performed in the Gingras laboratory by the two full-time trainees, Amber Couzens and Amanda Veri. The work in the Kim lab will be restricted to computation.

a. Reagents for mass spectrometry. Though an important part of the grant will be to streamline our procedures and reduce overall cost per experiment, the cost of materials for the proteomics experiments we are proposing is relatively high. All our experiments require sequencing-grade enzymes such as trypsin (~\$500). For the experiments using quantitative proteomics on the 5600 TripleTOF, we require RP-Chips for sample loading and separation (10 per year, at a cost of \$500/chip) Lastly, we require capillary tubing material, C18 reverse phase supports, HPLC plasticware, acetonitrile, HPLC-grade water and acids (formic acid, acetic acid, TFA) and bases (ammonium hydroxide), at an estimated cost of \$2,000.

\$7,500

b. Mammalian tissue culture. Last year, the cost of serum, selection medium and plates for the mammalian tissue culture experiments specifically for a project of similar size were ~\$ 4,000. For example, we have used ~6 cases of 150mm plates, 25 ml hygromycin, 4 bottles of FBS and 40 bottles of DMEM, not counting the transfection reagents and other supplies for which we spent another ~\$ 2,000. We anticipate similar costs for the coming years.

\$6,000

c. Affinity purification reagents. We require commercial antibodies and affinity reagents (e.g. large quantities of FLAG M2 magnetic beads) for our AP-MS experiments and custom antibodies for the other immunoprecipitation, immunoblotting and immunofluorescence experiments. This is a significant cost in our lab, and we expect to spend nearly \$6,000 yearly on affinity reagents related to this project.

\$6,000

d. Molecular biology and standard reagents. Inhibitors (e.g. protease inhibitor cocktail, phosphatase inhibitors) are estimated at \$1000. Oligonucleotides and cloning enzymes (including Gateway clonase and PCR enzymes) are required for cloning and mutagenesis (note that the project involves a large molecular biology component; \$2,700). Standard chemicals, molecular weight ladders and plasticware are also required for the project (\$2,500).

\$ 5,200

e. Publication expenses: Preparation of figures for poster presentations and publication costs including page charges and open access option for one paper each year.

\$ 1,500

YEARLY EXPENDABLES: \$26,200

4. TRAVEL

We request \$1,000 per year for one of the trainees (Couzens, Veri or Teyra) paid on this grant to partially cover the cost of travel to scientific meetings to present their research findings on this project. The Kim and Gingras labs are conveniently located within ~5

Gingras, Anne-Claude

Justification - supplies

minutes of each other by foot, enabling frequent exchanges, and common meetings as necessary. While the vast majority of our meetings with other international collaborators are via webinars, we also request \$500 each year for Amber Couzens to travel to the our collaborators' labs (year 1 = Vidal lab in Boston; year 2 = Nesvizhskii lab in Ann Arbor, Michigan) to facilitate the collaborations and learn new techniques.

\$ 1,500

YEARLY TRAVEL: \$ 1,500

Gingras, Anne-Claude

Justification - salaries

Budget justification:

A) Research Personnel:

Amber Couzens, Postdoctoral Fellow, 100% effort, 100% salary support requested from CCSRI

Supervised by Anne-Claude Gingras

Amber joined our group in 2011. She obtained her PhD in 2011 from York University for her studies on the regulation of ROCK kinase, and on the development of biochemical and proteomics approaches to identify substrates for this enzyme. She is very skilled at biochemical and mass spectrometry approaches, and has already generated interaction maps for ~30 proteins. She also optimized our pipeline for affinity purification coupled to mass spectrometry (she is second author on the manuscript detailing our optimized protocols; Kean et al., *Methods*, 2011). She has been involved in a large international effort to generate a Contaminant Repository for Affinity Purification (she will be an author on the manuscript which will be submitted shortly), and has been involved in several collaborative projects. Importantly, she has direct experience with the 5600 TripleTOF mass spectrometer and in particular with data independent acquisition, and she is the second author on the manuscript which will be submitted shortly on the AP-SWATH approach. Within the context of this project, she will direct the AP-SWATH experimental process, and perform the bulk of the experiments in Aim 2. For Aim 3, after identification of those proteins for which pharmacological compounds already exist, she will guide the experiments to analyze the effect of treatment with these compounds on protein-protein interactions (if the compounds target proteins identified in Aim 2, she will perform the experiments herself).

Amber will be in her third year in 2013, and I am requesting the annual maximum eligible from the CCSRI, namely \$38,165. (To be in line with \$42,000 plus benefits (15%) as mandated by the Samuel Lunenfeld Research Institute salary scale, or \$48,300, I will supplement Amber's salary directly from the funds allocated through my Canada Research Chair).

Amanda Veri, Graduate Student, 100% effort, 100% salary support requested from CCSRI

Supervised by Anne-Claude Gingras

Amanda Veri was a coop student in my group from May 2011 – Dec 2011 and a summer student from May 2012 – September 2012. During this time, she worked on a large scale interactome project involving all human phosphatases (~150 genes). Amanda distinguished herself by her positive attitude, her hard work, and her ability to accomplish difficult tasks and troubleshoot problems. She has already performed a large number of constructions, has established many stable cell lines, and performed LC-MS/MS analysis and data analysis. Therefore, though a relatively young student still, she is fully capable of performing the experiments in a timely manner. Within the context of this grant, she will be responsible for the experimental aspects of Aim 1 (namely to transfer the selected alleles into our expression vectors, generate stable cell lines, perform affinity purification, and SWATH mass spectrometry).

She will also work together with Amber Couzens on drug-regulated interactions in Aim 3.

Amanda's salary is \$25,402 as mandated by the school of Graduate Studies at the University of Toronto. I am requesting the maximum amount available through the

Gingras, Anne-Claude

Justification - salaries

CCSRI for a student stipend, \$19,500, and will cover the remainder of a stipend through my Canada Research Chair allocation.

Joan Teyra, Postdoctoral Fellow, year 2, 40% effort, 40% salary support requested from CCSRI

Supervised by Philip Kim

Joan is ideally suited to lead the structural analysis portion of the project. He has a BSc degree in biology, but a MSc in Bioinformatics (Madrid, Spain) and a PhD in Structural Bioinformatics (Dresden, Germany). In his PhD work, he has specifically worked on the large-scale analysis of protein-protein interactions (Teyra et al., *BMC Bioinformatics* 2011 and Teyra et al. *Proteins* 2011); he thus has ample expertise highly relevant for this application. Joan has been in the Kim lab for roughly a year and has been highly successful in his main project and has made impressive progress, such that he has several manuscripts in preparation. His ongoing research is highly complementary to this application – he is aiming to identify protein-protein interactions that are susceptible to inhibition and develop peptide-based compounds to inhibit them.

For this application, Joan will utilize structural methods to analyze the current collection of clones available to us for mutations most likely to disrupt protein-protein interactions, and help with the final selection of the proteins and mutants to be tested based on the criteria described in the application (Aim 1). He will also periodically revise his prediction tools based on the quantitative proteomics data generated in this project (in Aim 3), and help identifying drug targets within our interaction networks.

Joan is a second year postdoctoral fellow (eligible salary support from CCSRI = \$36,580 / year). We are asking only for the portion of his salary directly applicable to this CCSRI project, which we estimate at 40%, \$14,632. The rest of his salary is covered by Dr. Kim's other funding.

Amber and Amanda will perform the bulk of the mass spectrometry experiments proposed here with the point mutants and follow-up validation; Joan Teyra will perform computational analysis of the structural impacts of protein-protein interactions.

They will be aided for the mass spectrometry portion of her project by the Lunenfeld mass spectrometry coordinator, Brett Larsen, by postdoctoral fellow Jean-Philippe Lambert and by the Lunenfeld's proteomics data management and software developer, Guomin (Frank) Liu, for whom no salary is sought. Their respective roles are outlined below:

Brett Larsen (supervised by AC Gingras) is our most senior mass spectrometry expert at the Lunenfeld, and is in charge of maintaining the 12 mass spectrometers currently in operation (he is assisted in this by excellent technicians), training the students and postdoctoral fellows and developing technically challenging protocols for the analysis of samples by mass spectrometry, including optimization of the SWATH data-independent acquisition protocols. He will continue assisting Amber and Amanda on the cancer interactome project, while working with AC Gingras and our computational / industry collaborators on refining the SWATH acquisition and data analysis methods. Brett's salary is covered partly by the Samuel Lunenfeld Research Institute, and another portion of his salary, which will be used specifically for the development of AP-SWATH, has been applied to as part of a supplement to the NIH R01 to Nesvizhskii and Gingras.

Gingras, Anne-Claude

Justification - salaries

Frank (Guomin) Liu is the main programmer behind the LIMS system for interaction proteomics, ProHits, and is first author on the two ProHits manuscripts (Liu et al., *Nature Biotech*, 2010; Liu et al., *Current Protocols in Bioinformatics*, 2012). Frank is essential to the continued development of the software tools, including the implementation of quantification modules to incorporate MS/MS SWATH quantification which will be required for this project. Importantly, Frank already collaborates extensively with the Nesvizhskii group, which will facilitate implementation of the computational tools (for example he is second author on a recent manuscript by Choi et al. describing the use of the statistical tool SAINT for quantitative proteomics. Salary support for Frank Liu is being sought in the context of a Genome Canada Bioinformatics competition: in particular, the GC bioinformatics grant will enable us to streamline the data analysis part of the SWATH pipeline.

Jean-Philippe Lambert is a postdoctoral fellow co-supervised by AC Gingras and T Pawson who has already published 25 papers, many of them as a first author. Importantly, Jean-Philippe's PhD was in interaction proteomics (he was a graduate student with D Figeys in Ottawa), and he brought a strong expertise to our research groups. Jean-Philippe has been the driving force in the development of the AP-SWATH approach and is the first author on the manuscript to be submitted. Jean-Philippe's main research interests reside in the characterization of the acetylome machinery, but he fully realizes that optimizing the AP-SWATH pipeline is key to the realization of his own research objectives. In particular, he is exploring the consequences of treating cells with specific bromodomain inhibitors (such as the molecule JQ1) on the interactome established by bromodomain-containing proteins. As such, he will continue to work closely with Amber and Amanda on all technological aspects of the research project. Jean-Philippe holds a competitive CIHR postdoctoral award: the top-up to his salary is provided through a CIHR grant to Gingras, Filippakopoulos and Pawson on the Acetylome.

No salary is requested for personnel in the laboratories of our collaborators.

YEARLY SALARIES: \$ 72,297

Current Operating Funds

GINGRAS

Title of Grant: Global approaches to unravel PP2A function

Source: Canadian Institutes for Health Research Operating Grant

Dates of Approved Project: 2010-2015 **Term:** 5 years

Dollars requested: \$ 192077/year

Name of PI: Dr. Anne-Claude Gingras

% of effort for applicant on this grant: 15

List of Co-Applicants:

% of effort for those on this grant: N/A

The major goals of this project are: This application aims to improve our proteomics analytical pipeline (interactomics, phosphoproteomics), and to apply it to further our understanding of protein complexes surrounding the PP2A serine/threonine phosphatase. We will also provide further biological insight into the role of a newly identified protein complex, STRIPAK, in cerebral cavernous malformations.

% of Overlap: 0%. There is no overlap: we have used this grant to build a standard interaction proteomics pipeline that we use for static interaction studies. The work on PP2A focuses on its role in cerebral cavernous malformations (not cancer).

Title of Grant: Structure, function and regulation of PP4cs

Source: Canadian Cancer Society Research Institute (CCSRI)

Dates of Approved Project: 2009-2014 **Term:** 5 years

Dollars requested: \$137721/year

Name of PI: Dr. Anne-Claude Gingras

% of effort for applicant on this grant: 15

List of Co-Applicants:

% of effort for those on this grant: N/A

The major goals of this project are: This proposal aims at characterizing the role of the protein phosphatase 4 (PP4) in mechanisms such as transcriptional elongation and splicing. Structural analysis of a PP4 trimeric complex which was discovered by us are also proposed.

% of Overlap: 0%. There is no overlap: this grant is on protein phosphatase 4, which is not known to be mutated in cancers

Title of Grant: Understanding the assembly and function of dynamic signalling networks in complex diseases

Source: Ontario Research Fund – Global Leadership Round in Genomics

Dates of Approved Project: 2010-2015 **Term:** 5 years

Dollars requested: \$ 179634/ year

Name of PI: Dr. A.J. Pawson

% of effort for applicant on this grant: 10

List of Co-Applicants: Dr. Anne-Claude Gingras **% of effort for those on this grant:** 5-20

The major goals of this project are: This proposal seeks to understand how signalling pathways assemble into complex networks that regulate cellular behaviour and how these networks are altered in human disease. The research will help Ontario maintain leadership in the field of Cell Signalling and Systems Biology, train and retain HQP and yield new discoveries with commercialization potential.

% of Overlap: 0%. The project is only aimed at characterizing interaction network for wild type proteins. The technological platform developed within the context of the GL2 application was instrumental in our adoption of the SWATH platform, which is necessary for the current work. The main contribution of Dr. Gingras to this grant is technology development for mass spectrometry (the salaries supported by the portion of this grant to Dr. Gingras are for the key personnel in the core proteomics laboratory).

Title of Grant: Computational tools for mass spectrometry-based Interactome

Source: National Institute of Health (NIH) – RO1

Dates of Approved Project: 2010-2015

Term: 5 years

Dollars requested: \$30000/year to AC Gingras

Name of PI: Dr. Alexey Nesvizhskii

% of effort for applicant on this grant: 5

List of Co-Applicants: Dr. Anne-Claude Gingras

% of effort for those on this grant: 25

The major goals of this project are: This proposal aims at generating computational tools that help us extract protein-protein interaction information from mass spectrometry data.

% of Overlap: 0%. This project led to the development of the SAINT series of algorithms which we use for our static interaction maps. We are still collaborating with Dr. Nesvizhskii to extend the SAINT tools to the SWATH data, which will be complementary to this application. The funds to Dr. Gingras are for the core proteomics laboratory.

Title of Grant: Molecular Mechanisms of Cerebral Cavernous Malformations

Source: Canadian Institutes for Health Research Operating Grant

Dates of Approved Project: August 2012 – July 2017 **Term:** 5 years

Dollars requested: 105000\$/year to AC Gingras (total 190026\$/year)

Name of PI: Dr. Anne-Claude Gingras

% of effort for applicant on this grant: 15

List of Co-Applicants: Dr. Brent Derry

% of effort for those on this grant: 10

The major goals of this project are: In this proposal, we are combining genetics (Derry) and proteomics (Gingras) approaches to understand the signaling pathway(s) in which the three proteins mutated in cerebral cavernous malformations reside.

% of Overlap: 0%. No known link to cancer.

Title of Grant: A systems approach towards the therapeutic modulation of the acetylome

Source: Canadian Institutes for Health Research Operating Grant

Dates of Approved Project: Oct 2012-Sept 2016 **Term:** 4 years

Dollars requested: \$120000/year to AC Gingras (total 202115\$/year)

Name of PI: Dr. Anne-Claude Gingras

% of effort for applicant on this grant: 10

List of Co-Applicants: Dr. Panagis Filippakopoulos and Dr. A.J. Pawson

% of effort for those on this grant: 5-10

The major goals of this project are: In this application, we propose to systematically define the specificity and function of human bromodomains through identification of their binding partners and acetylated targets. This knowledge will assist in the development of new inhibitors of bromodomains, accelerating the development of new therapeutics.

% of Overlap: Financial overlap = 0%. This is largely an analysis of the binding site specificity of the bromodomain. There is some conceptual overlap with Aim 3 of the current application as we are testing within the context of this application the effect of small molecule inhibitors of the bromodomain on the protein-protein interactions established by the bromodomain-containing proteins. Please note however, that this is all done within the context of wild type proteins with the exception of translocation products; there is therefore no financial overlap with the current application.

Funds Requested

Title of Grant: ProHits 2.0: A flexible system for tracking, analyzing and reporting functional proteomics data.

Source: Genome Canada

Dates of Approved Project:

Term: 2 years

Dollars requested: \$ 250,000/year for the entire team

Name of PI: Dr. Anne-Claude Gingras

% of effort for applicant on this grant: 5

List of Co-Applicants: Dr. Michael Tyers, Dr. Pierre Thibault, Dr. Alexey

Nesvizhskii, Dr. Tony Pawson and Dr. Hyungwon Choi.

% of effort for those on this grant: 5-10

The major goals of this project are: In this proposal, we aim at refining our software tools to better distribute them to the research community.

% of Overlap: 0%. This grant application is solely computational and aims at the integration of computational tools generated across the laboratories of all co-applicants. The funds assigned to the Lunenfeld are for the core proteomics laboratory.

Title of Grant: Personalized signalling signatures to guide therapy in cancer

Source: Genome Canada

Dates of Approved Project:

Term: 4 years

Dollars requested: \$10,188,446 (no direct funding to the Gingras lab – the involvement of Dr. Gingras is through her role as co-director of the mass spectrometry facility)

Name of PI: Dr. AJ. Pawson

% of effort for applicant on this grant: 5

List of Co-Applicants: Dr. Robert Rottapel (co-PI), Dr. Amit Oza, Dr. Anne-Claude Gingras, Dr. Gary Johnson, Dr. Olga Vitek, Dr. Benjamin Neel, Dr. James Woodgett, Dr. Murray Krahn, Dr. Michael Roehrl, etc.

% of effort for those on this grant: 5-15

The major goals of this project are: This proposal aims at using signaling readouts to better predict the response to therapy of patients with ovarian cancer.

% of Overlap: 0%.

KIM

Title of Grant: Large-scale development of inhibitors to intracellular cancer targets using an integrated computational and combinatorial approach

Source: Canada Institute for Health Research, Operating Grant

Dates of Approved Project: 2012-2017

Term: 5 years

Dollars requested: \$170990/year

Name of PI: Dr. Philip Kim

% of effort for applicant on this grant: 20

List of Co-Applicants: N/A

% of effort for those on this grant:

The major goals of this project are: Develop Novel Cancer Inhibitors

% of Overlap: 10

Title of Grant: Development of highly diverse but fully defined phage display libraries

Source: Ontario Genomics Institute

Dates of Approved Project: 2012

Term: 6 months

Dollars requested: \$ 50,000

Name of PI: Dr. Sachdev Sidhu

% of effort for applicant on this grant: 5

List of Co-Applicants: Dr. Phillip Kim

% of effort for those on this grant: 5

The major goals of this project are: Develop new technologies for phage display

% of Overlap: 0

Title of Grant: Synthetic antibody program: commercial reagents and novel therapeutics

Source: Genome Canada

Dates of Approved Project: 2011-2013

Term: 3 years

Dollars requested: \$10,000,000

Name of PI: Dr. Sachdev Sidhu

% of effort for applicant on this grant: 5

List of Co-Applicants: Dr. Philip Kim, Dr. Charles Boone, Dr. John Dick, Dr. Tony Pawson, et al.

% of effort for those on this grant: 10

The major goals of this project are: Develop novel antibody therapeutics

% of Overlap: 0

Title of Grant: Analysis of Evolutionary Mechanisms in Signaling Networks

Source: Natural Science and Engineering Research Council

Dates of Approved Project: 2010-2016

Term: 6 years

Dollars requested: \$122,000

Name of PI: Dr. Philip Kim

% of effort for applicant on this grant: 10

List of Co-Applicants: N/A

% of effort for those on this grant:

The major goals of this project are: Elucidate mechanisms of network evolution

% of Overlap: 0

Title of Grant: Leadership Opportunity Fund

Source: Canada Foundation for Innovation

Dates of Approved Project: 2010-2015

Term: 5 years

Dollars requested: \$ 233,000

Name of PI: Dr. Philip Kim

% of effort for applicant on this grant: 10

List of Co-Applicants: N/A

% of effort for those on this grant:

The major goals of this project are: Built a structural human interaction network

% of Overlap: 10. Note that this is an infrastructure grant only: the resources acquired within the context of this application are critical for the realization of the structural predictions proposed here.

GINGRAS, Anne-Claude

Operating Grant/Subvention de fonctionnement Application/Demande 2010-03-01

Summary of Research Proposal/Résumé de la proposition de recherche

Phosphorylation plays critical roles in nearly all cellular processes. The serine/threonine phosphatase PP2A accounts for nearly half of the serine/threonine phosphatase activity in the cell, and its deregulation has been linked to multiple human cancers (where PP2A acts as a tumor suppressor), Alzheimer's disease, and increased susceptibility to pathogen infection. The catalytic subunit of PP2A acquires specificity in part via association with adapter molecules, which confer intracellular localization, activation and/or substrate recognition cues. How each of the different PP2A-containing holoenzymes is regulated and directed to its substrates still remains largely uncharacterized.

We have applied iterative affinity-purification coupled to mass spectrometry (AP-MS) to characterize PP2A-containing protein complexes. While the most abundant form of PP2A is a trimer consisting of the catalytic (C) enzyme associated with a scaffolding (A) subunit and one of many regulatory (B) subunits, our recent work also identified higher order stable complexes. One of these complexes, which we termed STRIPAK (for STRiatin Interacting Phosphatase and Kinase), contains at its core a PP2A ABC trimer, in which the B subunit is striatin. Associated with this core structure are several uncharacterized proteins, as well as kinases of the Ste20 family and the protein CCM3. Interestingly, the CCM3 (PDCD10) gene is mutated in familial cases of cerebral cavernous malformations (CCMs), vascular anomalies of the central nervous system in which the structure of the blood vessel lacks integrity, and characterized by accumulation of stagnant blood in "caverns", along with sporadic bleeding. Familial CCM is a monogenic disease that can result from the mutation of either CCM1, CCM2 or CCM3; mutation results in the loss of protein expression, and the disease is transmitted in an autosomal-dominant fashion, implicating a two-hit mechanism. While CCM1 and CCM2 are part of the same protein complex, CCM3 largely resides within STRIPAK. CCM3 lesions are more virulent than CCM1/2 mutations, and the lack of understanding of these proteins drastically limits treatment options (currently, surgery is the principal option).

This application aims to improve our proteomics analytical pipeline (interactomics, phosphoproteomics), and to apply it to further our understanding of protein complexes surrounding the PP2A serine/threonine phosphatase. We will also provide further biological insight into the role of a newly identified protein complex, STRIPAK, in cerebral cavernous malformations.

Aim 1. Develop robust tools for mass spectrometry-based interactomics, and apply them to the study of differential PP2A interaction networks.

Aim 2. Develop and deploy a phosphorylation analysis pipeline for the identification of pathways and substrates of PP2A complexes.

Aim 3. Define the molecular mechanism of action of STRIPAK, a PP2A-containing complex with links to cerebral cavernous malformations.

We expect three major outcomes of our studies: 1) development of improved toolboxes for proteomics that will be widely distributed to the research community in Canada and elsewhere; 2) a better understanding of a multifunctional protein phosphatase; 3) uncovering avenues of investigation for treating cerebral cavernous malformations.

CCSRI grant**ITEM 10****Detailed scientific abstract**

Gingras, Anne-Claude (PI)

Samuel Lunenfeld Research Institute; Dept Molecular Genetics, U of Toronto

Title of project: Structure, regulation and function of the cisplatin-sensitive protein phosphatase 4 complex

Keywords/technical terms: (5-10) cisplatin, signal transduction, PP4, phosphatase, mass spectrometry, protein complex, affinity-purification.

Protein phosphatase 4 is a serine/threonine phosphatase related to protein phosphatase 2A (PP2A), yet exhibiting non-overlapping functions. In addition to intrinsic specificity of the catalytic subunits, PP2A-related phosphatases acquire specificity via association with a variety of different regulatory subunits, which provide substrate targeting, subcellular localization or activation cues. Our laboratory has been interested in identifying and characterizing these regulatory subunits. We previously reported the identification of an evolutionarily conserved complex containing the catalytic subunit of PP4 (PP4c), a previously-known regulatory subunit (PP4R2), and a novel subunit which we termed PP4R3 (two highly related proteins are present in vertebrates). In yeast, all three components have been implicated in sensitivity to the anticancer drug cisplatin, and dephosphorylation of the histone variant H2AX during recovery from DNA damage. We recently demonstrated that human PP4c is also a phosphatase for H2AX in human cells, and that it contribute to recovery from the DNA damage checkpoint.

We have developed high-sensitivity affinity-purification coupled to mass spectrometry (AP-MS) to investigate the cellular context for PP4 components. The generation of high-density interaction maps around each of the subunits allows us to form additional hypotheses regarding the function of PP4 in human cells. We have identified many novel proteins as interactors for the PP4c•PP4R2•PP4R3 complex that largely fall into two categories: 1) splicing components; 2) transcription elongation factors. Both of these processes are critical for proper cell function, and are misregulated in cancer. Both processes are also regulated by phosphorylation, although the phosphatase(s) responsible have not been identified. Furthermore, many of the components identified here exhibit dual functions (e.g. the splicing factor hPrp19 is also known as PSO4 and is involved in DNA damage repair). Lastly, one remaining interactor, PTC1 (papillary thyroid carcinoma gene 1), is a frequent fusion partner for the RET oncogene in thyroid cancers that has been linked to DNA damage. Our preliminary data indicates that the phosphorylation of PTC1 is altered by co-expression of PP4c, indicating that it may constitute a substrate for the enzyme.

We propose to further decipher the roles on PP4c on these putative substrates and cellular process by performing the following aims:

Aim 1: Further characterization of the PP4 physical interaction network, and phosphoregulation by specific PP4 complexes

Aim 2: Defining the role of PP4 in DNA damage, splicing and transcription elongation

Aim 3: Structural studies of PP4 complexes

6.b RESEARCHERS INTEREST IN PRIVATE SECTOR PARTNER(S) If applicable, indicate if any researcher involved in the project has any interest in any of the private sector partners named in the proposal. Describe the extent of the interest, i.e. the relationship between the researcher and the company (see Private Sector Partners section in guidelines)

Our researchers have no interest in private sector partners.

7. LEAD INSTITUTION CONTACT	
Name:	Dr. Gareth Taylor
Title:	Director, GRIP
Institution:	Mount Sinai Hospital
Department:	Samuel Lunenfeld Research Institute
Courier Address:	60 Murray Street, Room L1-020, Toronto, Ontario, M5T 3L9
Telephone:	416-586-4800 X 3125
Fax:	416-586-4877
E-mail :	grtaylor@lunenfeld.ca

7a. PRINCIPAL INVESTIGATOR	
Name:	Dr. Tony Pawson
Title:	Distinguished Investigator
Institution:	Mount Sinai Hospital
Department:	Centre for Systems Biology, Samuel Lunenfeld Research Institute
Courier Address:	600 University Avenue, Room 1081, Toronto, Ontario, M5G 1X5
Telephone:	416-586-4800 X 8262
Fax:	416-586-8869
E-mail :	pawson@lunenfeld.ca
Overall number of co-investigators:	10

Quality of Research

8a. PROPOSAL BRIEF

(Describe in no more than three lines, **using simple language**, the objective and key elements of this proposal, including the strategic value to Ontario. **Please Note:** This summary description may be used, in whole or in part, in press releases or similar materials, if the project is approved.)

This proposal seeks to understand how signalling pathways assemble into complex networks that regulate cellular behaviour and how these networks are altered in human disease. The research will help Ontario maintain leadership in the field of Cell Signalling and Systems Biology, train and retain HQP and yield new discoveries with commercialization potential.

8b. PROPOSAL ABSTRACT

(Provide a proposal abstract of no more than 500 words, **using non-technical language**. **Please Note:** Do not include images and/or charts in the summary description. It may be used in press releases or similar material, if the project is approved.)

Coordination of cell behaviour by signalling networks is a fundamental mechanism underlying all animal development, from simple worms to complex mammals. Genome sequencing led to the surprising revelation that increasing complexity is not reflected in increased numbers of genes. This is consistent with the parallel revelation that the development of animals is governed by a small set of signalling pathways that are used over and over to direct the specification of diverse cell types and their intricate arrangement into organs. The reason that this simple strategy can give rise to the intricate patterns that confer function to organs is that the proteins are dynamically organized into vast networks. This provides for contextual response that can direct the divergent outcomes required for tissue morphogenesis. Moreover, these systems frequently go awry in complex diseases associated with aging such as cancer, where the resulting pathological changes often lead to metastasis and death. For example, we recently showed that changes in gene expression that alter the global topology of the human interactome are associated with poor outcome breast cancers. Therefore, the **main objectives of this proposal are to understand how protein interaction networks (interactomes) are assembled and how their topology and function are modulated by post-translational modifications, gene expression and alternative splicing.**

To achieve these objectives a multidimensional approach to mapping interactomes will be taken. We will integrate mass spectrometry-based approaches (Pawson, Gingras, and Aebersold) with domain-based approaches (Li) and high throughput

screens (Wrana, Vidal) to generate focused maps of the human interactome. To understand the dynamic regulation of cell function and protein interactions, we will focus on protein kinases and phosphatases, the versatile properties of scaffold and adaptor proteins, the regulation of Rho GTPases, and the control of cell polarity (Pawson, Gingras, Wrana, Aebersold, Li). Despite the fact that complex splicing events provide enormous diversity to the human proteome, very little is known of how this diversity controls interactome dynamics. Therefore, in a second major direction, we will exploit deep sequencing of RNA (RNA-seq) in normal and cancer cells and tissues (Blencowe, Morris) to map splicing diversity. This will be coupled to splice-specific mass spectrometry tools, affinity reagents, high throughput interactome mapping and novel panels of splice-specific siRNAs with the goal of understanding how splicing affects network dynamics and cell function (Blencowe, Gingras, Pawson, Wrana). In addition, RNA-seq in different cell subtypes in solid tumours will allow us to explore the complex interplay between tumour cells and the stroma that is a critical driving force in tumour progression. Finally, the group will develop novel computational tools to analyze how the major mechanisms of generating interactome dynamics provide for contextual responses in both normal and pathological states.

This project will be integrated with local and international commercial partners to develop new technologies in the area of proteomics that can be directly commercialized through our industrial partners. In addition, we expect to derive new network-based biomarkers for cancer and define new types of targets for the design of therapeutics, for example splice-specific targets.

8c. PROPOSAL DESCRIPTION

(Describe your research plan and scientific methodology with a maximum of 15 pages for projects under \$5 million (for the ORF request) and a maximum of 25 pages for projects over \$5 million (for the ORF request). See Application Instructions for further detail.)

OVERVIEW

Coordination of cell behaviour by signalling networks is a fundamental mechanism underlying all animal development, from simple worms to complex mammals. Genome sequencing led to the surprising revelation that increasing complexity is not reflected in increased numbers of genes. This is consistent with the parallel revelation that the development of animals is governed by a small set of signalling pathways that are used over and over to direct the specification of diverse cell types and their intricate arrangement into organs. Moreover, these systems are abused in complex diseases of aging such as cancer. The reason that this simple toolkit can give rise to the intricate patterns that confer function to organs is that the proteins are dynamically organized into vast networks. This provides for contextual response that can direct the divergent outcomes required for tissue morphogenesis, but in cancer drives the pathological changes that lead to metastasis. For example, we recently showed that changes in gene expression that alter the global topology of the human interactome are associated with poor outcome breast cancers. Therefore, the main objectives of this proposal are to understand how signalling pathways assemble into dynamic networks by post-translational modifications (PTMs), gene expression and alternative splicing.

INTRODUCTION

Signalling networks are intimately enmeshed within the human interactome and function in a dynamically regulated fashion to coordinate virtually all cellular responses to extrinsic and intrinsic cues. The flow of information through signalling pathways within the cell is thus a highly dynamic process that is normally exquisitely regulated with respect to the subcellular location at which signalling events take place, and the kinetics with which they are activated and suppressed. The molecular devices employed to control the transmission of signals from receptors to their intracellular targets include a range of post-translational modifications (PTM), for which protein phosphorylation is the prototype, as well as regulated protein-protein and protein-phospholipid interactions, targeted proteolysis, and the switch-like allosteric activation of key signalling proteins such as protein kinases and small GTPases¹. These various modes of regulation are used in a reiterated fashion in many different signalling pathways, and frequently act in coordination to establish the spatial and temporal organization of cellular responses to specific stimuli. For example, phosphorylation can induce a conformational change in enzymes such as protein kinases, and thereby promote a switch from an active to an inactive state. However, the most common effect of protein phosphorylation, be it on tyrosine or serine/threonine residues, is to form a binding site for phosphorylation-dependent protein interaction domains, such as SH2 domains in the case of phosphotyrosine or FHA domains for phosphothreonine (as two of many examples)². In this latter case, phosphorylation induces intra- or intermolecular protein-protein interactions that can control protein activity and the formation of multi-protein complexes essential for signal propagation. Modification-dependent protein-protein interactions can also function in series, as exemplified by the numerous links between phosphorylation and ubiquitination, where phosphorylation of a protein creates a binding site for the substrate recognition domain of an E3 protein-ubiquitin ligase, which consequently ubiquitinates the phosphorylated target^{3,4}. Such ubiquitinated sites then associate with the ubiquitin-binding domains of downstream proteins, leading to fates such as degradation by the proteasome, activation of signalling pathways involved in innate immunity, or trafficking to endosomes.

Studies of cell signalling have, for the most part, focused on specific pathways, and have provided us a rich and detailed view of how individual core signalling pathways regulate cellular behaviour. However, *in vivo*, cells are exposed to multiple signals that

Principal Investigator/Program Director (Last, First, Middle):

Nesvizhskii, Alexey I.

A. SPECIFIC AIMS

The analysis of protein complexes and protein-protein interaction (PPI) networks - and the dynamic behavior of these networks as a function of time and cell state – are of central importance in biological research. The recent technological advances have made affinity purification and mass spectrometry (AP/MS) a high-throughput technique. At the same time, the development of computational methods and tools for AP/MS data has lagged behind. There is a great need for accurate and robust methods for statistical assessment of PPI data. Furthermore, AP/MS data is presently used to reconstruct protein networks that are purely qualitative in nature. Such important questions as partition of proteins into multiple complexes, and changes in the complex composition upon perturbation in the cellular environment remain to be fully addressed. We propose to add a new dimension to MS-based analysis of PPI networks and complexes by taking advantage of the quantitative information encoded in MS data. We have demonstrated that AP/MS can be coupled with a label-free spectral counting approach for relative protein quantification. This can provide the basis for more accurate reconstruction of protein complexes and interaction networks. To this end, we will develop a novel statistical approach, Significance Analysis of Interactome (SAINT), which will utilize spectral counting for assigning a confidence measure to individual interactions. This work will reduce the burden of conducting numerous additional experiments for negative controls. It will also allow assessment of the reproducibility of interactions across multiple biological replicates or affinity tags. The proposed methods will be applied in two key areas of biological research linked through their significance for fundamental understanding of cell signaling: 1) identification of the components of protein complexes involving the serine/threonine protein phosphatases in human cells and the effects of viral infection on these complexes; 2) large-scale analysis of protein kinases and their interactions in budding yeast. We will integrate the new methods with the open source data analysis system ProHits, and will disseminate the new tools, statistical methodologies and educational materials to the research community. The ultimate goal of the proposed computational research is to enable generation of high quality PPI networks and complexes from AP/MS data and their subsequent biological interpretation. Our specific aims are:

Aim 1. To develop a robust computational framework for validation of protein interaction data We will develop a novel framework (SAINT) for assessing the significance of protein interactions using normalized spectral counts as a measure of protein abundance in the affinity purified sample. We will start by creating a database of non-specific proteins observed in different control AP/MS experiments. We will then develop a method for statistical validation of PPI by probabilistic modeling of protein spectral count profiles across multiple AP/MS runs. Finally, a combined approach that effectively marries these two strategies via semi-supervised probabilistic modeling will be developed. The method will be extensively tested using gold standards datasets created with the help of stable isotope labeling based quantitative proteomics.

Aim 2. To define the effect of viral tumor antigen expression on PP2A interactome We will use the high density PPI network centered around the protein phosphatase 2A (PP2A) generated in our previous studies as a starting point for the development of methods for quantitative analysis of protein networks. We will investigate how changes in the PP2A network caused by viral proteins can be effectively interrogated using AP/MS, SAINT analysis, and novel model-based clustering approaches that can utilize label-free quantitative information. We will also develop computational method enabling more effective interactome mapping.

Aim 3. To develop computational methods for systematic interrogation of large-scale PPI networks In collaboration with Mike Tyers, we will apply our computational methods to perform systematic interrogation of the budding yeast kinome data. SAINT will be used to select high confidence PPI from more than 1500 AP/MS experiments covering all 131 kinases in budding yeast. A rigorous statistical model will be developed for assessing the reproducibility of identified PPI over replicate experiments. Network modeling will be applied to functionally characterize sub-graphs and hubs within the kinase network, and to uncover distinct biological modules and functions connected by these hubs.

Aim 4. To disseminate computational tools and data analysis guidelines to the research community All tools, raw MS data, data annotation standards and analysis guidelines, acquired or developed as a part of this work, will be distributed freely and as an open source repository?. All newly developed computational methods, including the common contaminant data repository, will be integrated with the ProHits data management system.

GINGRAS, Anne-Claude

Operating Grant/Subvention de fonctionnement Application/Demande 2012-03-01

Summary of Research Proposal/Résumé de la proposition de recherche

Lysine acetylation on histones is key to the epigenetic regulation of gene expression, and is mediated by the action of acetyltransferases (writers) and deacetylases (erasers). Not surprisingly, the enzymes implicated in such a critical function are often found altered in cancers; in this regard, histone deacetylase inhibitors now constitute a prime target for cancer therapies. It has been recently realized, however, that acetyltransferases and deacetylases modify not only histones, but also a wide variety of proteins throughout the cell. Like phosphorylation, which can be recognized by dedicated protein modules (or “readers”, e.g. the SH2 domain), acetylation is recognized by a specific protein domain, the bromodomain. Bromodomain-containing proteins are often mutated and/or amplified in cancers, and bromodomains, which have deep pockets to accommodate acetylated lysine residues, are also viable targets for small molecule inhibition (for example, the BRD4 inhibitor JQ1 is a potent antitumor agent). Unfortunately, our current knowledge of bromodomain specificity and function is limited, hampering the rational design of therapeutic agents.

The Gingras and Pawson laboratories have developed extensive expertise in interaction proteomics and in the identification of post-translational modifications. Within the context of this application, we have recently developed optimized methods to identify protein-protein interactions for proteins associated with chromatin and identify and accurately quantify the acetylated sites. We have already cloned and stably expressed in mammalian cell lines 55 of the 75 proteins known to participate in acetyl lysine signalling, and initiated a systematic interaction mapping effort. This work has been done in coordination with a systematic effort at the Structural Genomics Consortium (SGC; Oxford) by co-applicant Filippakopoulos to systematically determine the structures of all human bromodomains and begin to define their specificity for acetylated lysines *in vitro* (Appendix 1; Filippakopoulos et al., *Cell*, *in press*). Through the efforts at the SGC, molecular probes that target the interaction of the bromodomains with acetylated peptides have also been developed. In summary, in the past year we have built the technical know-how and a series of resources which put us in an excellent position to provide a new view of the acetylome.

Here, we propose to systematically define the specificity and function of human bromodomains through identification of their binding partners and acetylated targets. This knowledge will assist in the development of new inhibitors of bromodomains, accelerating the development of new therapeutics.

We propose the following specific aims:

- 1) Identify the interactome for all components of the human acetylation machinery, including translocation products and mutated proteins
- 2) Define the specificity for each bromodomain-containing protein towards acetylated lysines
- 3) Reconstitute the acetylome specificity map
- 4) Test new bromodomain inhibitors

This project will enable us to better understand the function for each bromodomain-containing protein, as well as the mode of recognition of acetylated lysine residues by the bromodomain. With the growing interest in developing specific bromodomain inhibitors, this knowledge and the reagents generated will serve to rapidly assess, in a cellular context, the consequences of treatment with bromodomain inhibitors, and accelerate their progression to preclinical studies.

GINGRAS, Anne-Claude

Operating Grant/Subvention de fonctionnement Application/Demande 2012-03-01

Summary of Research Proposal/Résumé de la proposition de recherche

Cerebral cavernous malformations (CCMs) are common lesions of the brain that arise from a weakening of the microvasculature and result in a variety of pathophysiological conditions, including headaches, seizures and stroke. There is currently no cure or pharmacological treatment options for this disease, though surgical resection is available in severe cases to alleviate symptoms. CCM can arise by inherited mutations in one of three genes (CCM1, CCM2 or CCM3) and sporadically by unknown genetic loci. In familial CCM, loss of function mutations in CCM1, CCM2 or CCM3 account for at least 80% of cases, where the mutated allele is transmitted in an autosomal dominant fashion and loss of heterozygosity occurs in a subset of endothelial cells leading to the disease phenotype. While mutations in any one of these three genes results in similar histological manifestations, mutations in CCM3 are associated with the highest risk of hemorrhage. Patients with the same mutations in either one of the familial CCM genes have a range of symptoms and severity, suggesting the existence of modifier genes. Thus, understanding the biochemistry of CCM proteins and how these proteins engage their signaling pathways should not only uncover modifiers of familial disease but possibly genes that predispose to sporadic CCMs. To understand the *in vitro* and *in vivo* functions of the CCM proteins we propose to combine the powerful and complementary methods of proteomics and genetics. Biochemically, CCM1 and CCM2 associate with each other in a tight complex to control common downstream signaling events that modulate cytoskeletal function. Genetically, reports in mouse, zebrafish, and our preliminary data in *C. elegans*, indicate that CCM3 exhibits functions that are largely distinct from CCM1 and CCM2. While CCM3 can physically associate with CCM1•CCM2, results from our group and others suggest that most CCM3 protein resides in different molecular complexes in various cell types, including the newly described STRIPAK (STRiatin Interacting Phosphatase And Kinase) complex. Within STRIPAK, CCM3 is responsible for bridging kinases of the Sterile-20 family (GCKIII subfamily) to the striatin molecule, itself a regulatory subunit for the PP2A family of phosphatases. Interaction of CCM3 with the CCM1•CCM2 complex is also detected, but with lower stoichiometries than the CCM3•STRIPAK complex. Importantly, however, the GCKIII kinases were also detected by proteomics approaches in CCM1•CCM2•CCM3 complexes. Based on our functional characterization of roles for STRIPAK in cytoskeletal events such as regulation of Rho signaling and polarization, we postulate that the more severe phenotypes associated with CCM3 mutations in patients may be caused by concomitant loss of the CCM1•CCM2•CCM3 and STRIPAK•CCM3 signaling function. Consistent with this biochemical evidence, we have found that ablation of both *ccm1* and *ccm3* in *C. elegans* results in synthetic lethality, which we will exploit to screen for new components of the CCM3 pathway and the CCM regulatory network. In addition, we have found that *ccm3* and the STRIPAK complex are required for development the *C. elegans* excretory cell, a unicellular vascular structure that resembles capillaries of vertebrates.

The goals of this project are to uncover the molecular pathways that lead to CCM pathology using the powerful and complementary methods of proteomics and model organism genetics. To accomplish this we propose the following three Specific Aims:

Aim 1: Establish physical interaction maps for CCM proteins in mammals and *C. elegans*

Aim 2: Define the global genetic interaction map for CCM1 and CCM3

Aim 3: Identify substrates for GCKIII kinases

Taken together, these studies will not only greatly expand our understanding of CCM biology but potentially uncover new genes for preventive screening and the future development of therapeutics for treating patients with this disease.

VI RESEARCH SUMMARY

ProHits 2.0: A flexible system for tracking, analyzing and reporting functional proteomics data.

Understanding how proteins physically associate is key to uncovering their function(s). In recent years, affinity purification coupled to mass spectrometry (AP-MS) has become the approach of choice for mapping protein-protein interactions (PPIs). While the approach is efficient, tracking the large amount of data generated for systematic AP-MS projects alongside the experimental conditions is difficult, especially as some projects extends over months or even years. Additionally, AP-MS recovers a mixture of true interactors and background contaminants, and scoring methods that can discern enrichment of a protein in a purification of a bait as compared to a control are critical. Our team has an excellent track record of working together to address these important issues: we have developed the first LIMS for AP-MS, ProHits (Liu et al., *Nat Biotech*, 2010), and a widely used statistical tool for interaction scoring termed SAINT (Significance Analysis of INTeractome; Choi et al., *Nat Methods*, 2011). We have also created a reagent repository for high throughput collections (OpenFreezer; Olhovsky et al, *Nat Methods*, 2011). Lastly, we have been working on providing structural context to PPIs. Here, we propose the following main objectives that build on our collaborations:

- 1) Better tracking and visualization of quantitative data in ProHits. The current version of ProHits is based on a relatively crude quantification measure, namely, spectral counting. In recent years, we and others have however demonstrated that AP-MS can be coupled to more accurate quantitative approaches (such as intensity measurements; e.g. Bisson et al., *Nat Biotech*, 2011). In this aim, we will build new tracking and visualization tools to facilitate accurate quantification of AP-MS data, especially in the context of dynamically-regulated interactions.
- 2) Score, visualize and interpret PTM data within ProHits. Often, protein-protein interactions are modulated in a dynamic fashion by post-translational modifications. While ProHits has a rudimentary way to deal with PTM data, here we will build on the collective expertise of the group to provide functional annotation to PTM data (via the ProteoConnections software tool; Courcelles et al., *Proteomics*, 2011) and PTM site localization scoring.
- 3) Develop tools to provide structural information. We will develop and incorporate new bioinformatics tools for structural annotations of PPIs based on crosslinking studies which will provide a much higher resolution to our protein-protein interaction studies.
- 4) Setting the stage for data-independent acquisition (DIA). There is currently a paradigm shift in the mass spectrometry field to move from data-dependent acquisition (in which the n most abundant proteins are sequentially sequenced, leading to random sampling effects and incomplete datasets) to an unbiased DIA, and we have demonstrated that this approach is efficient to quantify interactome changes (Lambert et al., *in prep*). However, there is currently a dearth of tools to analyze DIA data. Here, we will further develop informatics tools for DIA and integrate them within ProHits.
- 5) Integration and dissemination. Collectively, we have already developed a number of complementary tools which will now be integrated to facilitate functional proteomics analysis. All tools will continue to be distributed freely as open source alongside documentation to facilitate development of new modules or workflows using the ProHits architecture.

VI RESEARCH SUMMARY

In a maximum of one page, summarize the proposed research, including GE³LS research activities. Describe the deliverables expected at the end of the project and the socio-economic benefits anticipated from their practical application.

Extensive research into signal transduction pathways that control cellular behaviour has led to the development of molecularly targeted therapeutics that directly attack the cellular proteins responsible for tumour growth and progression. Although targeted therapeutics have shown efficacy in the clinic, their success is often offset by limited patient response and the acquisition of resistance in the responders. To overcome these limitations and significantly impact future clinical care, it is necessary to gain a more comprehensive understanding of how signal transduction pathways differ in and between tumours and how drug treatment rewrites these pathways. While genetic and microarray analyses have provided tantalizing clues, real insight can only truly be gained by studying the actual proteins that comprise these signalling networks in cancers.

Focusing initially on women with advanced ovarian cancer (OC) who have relapsed from current standard treatment of platinum-based drugs, we propose to use quantitative mass spectrometry (MS) to create a laboratory-developed test (LDT) that will direct patients towards more effective second-line treatment and also function as an early surrogate assessment point for response once the therapy is administered. Currently, there are no targeted therapeutics that have received approval for OC in North America although several have shown promise in clinical trials. We will focus on signalling signatures downstream of receptor tyrosine kinases (RTKs) because there are a number of therapeutics directed to RTKs for which our LDT could be developed.

In this project, we will analyze alterations in multi-protein complexes that are assembled at cell surface receptors and the subsequent activation of downstream kinases to obtain distinct classes of response signatures across tumour sets. In **activity 1**, we will develop the signatures that will form the basis of our LDT starting with OC cell lines that we have extensively studied by functional genomics. In **activity 2**, we will optimize our MS technology for subsequent clinical application. Clinical trials for obtaining patient samples to build and validate the LDT will be conducted in **activity 3**. These signatures will assess patients prior to treatment and at two points after treatment (at 3-7 days which is the proposed surrogate endpoint for our test and at 56 days which is the standard clinical assessment time point). We expect that the biopsy prior to treatment will assess if the patient is likely to respond to treatment and the surrogate end point at 3-7 days will assess if response is actually occurring. As life expectancy in the highest risk patient population is short (6-9 months), improving precision of treatment selection and being able to determine efficacy or futility of therapy within a very short time will provide a more personalized and evidence-based approach to treatment of OC, and reduce time on ineffective therapy. It also ensures that patients who remain on therapy are most likely to benefit, which is important when considering issues of treatment toxicity and quality of life. In **activity 4** (GE³LS), we will use technology health analysis modelling to assess cost-effectiveness.

This approach will ultimately have broad applications in cancer care. We expect multiple benefits from this research including increased understanding of how OC treatments alter signalling pathways, and the development of signalling signatures that may prove effective as an LDT to guide treatment and improve therapeutic efficacy and potentially life expectancy. There will be intellectual property developed around the assays that will serve as a foundation for a Reference Laboratory that we will create to administer the proposed and future LDTs. A commercialization team will be established to ensure effective translation of the research. We anticipate directing patient care at project end.

Our team has extensive experience in all aspects of this project from signal transduction, mass spectrometry, conducting early phase clinical trials, clinical assay development and health technology assessment. We have engaged clinicians, regulatory experts, patient advocates and industry partners as end users for our technology.

Personalized signalling signatures to guide therapy in cancer

KIM, Philip

Operating Grant/Subvention de fonctionnement Application/Demande 2012-03-01

Summary of Research Proposal/Résumé de la proposition de recherche

Title: Large-scale development of inhibitors to intracellular cancer targets using an integrated computational and combinatorial approach

Background: There has been great recent progress in both wet-lab and computational methods to detect, perturb and design protein interactions. In particular, our knowledge of interactions mediated by short amino-acid stretches, so-called linear motifs, has grown tremendously. On the computational side, first-principle based structure modeling is providing useful insight for understanding and designing existing interactions and for simple predictions of new interactions that are close to a known template. Moreover, data driven machine-learning techniques are growing more powerful to predict novel interactions. On the side of bench technologies, modern methods such as high-throughput phage display or peptide arrays have revolutionized the way we study protein-peptide interactions. We will use a combination of the aforementioned approaches to develop novel intracellular inhibitors that, ultimately, may serve as templates for future drugs.

Rationale: Despite significant efforts of the whole biomedical research community, the number of target proteins for such small molecules has stagnated in recent years and remained largely confined to a very few protein families, most prominently the protein kinase family. Notable exceptions include the Bcl-2 family that has led to successful therapies. In this instructive case, first, a high affinity peptide was developed that ultimately led to the creation of a small molecule drug. We propose here to replicate this approach on a large scale and scan the human proteome for domains susceptible to peptide inhibitors, prioritize the putative targets according to biological and biochemical features and finally develop and validate inhibitors.

Specific Aims:

Aim 1: We will computationally identify proteins that are targetable using peptide inhibitors. We can identify potential peptide binders using two different techniques of different coverage and confidence. On the one hand, we will select a relatively small set of high confidence targets. These are those proteins that possess a domain that possesses a high level of sequence identity to known peptide binding domains. On the other hand, we will make use of recently established techniques to identify peptide-binding domains from known crystal structures of protein complexes. In short, this method is based on the idea that for a large subset of the structures, a short peptide contributes a large fraction of the binding energy and will thus also be a likely binder to this protein.

Aim 2: Prioritize proteins according to likely biological effects. We will then develop a prioritization scheme for these targets to focus on proteins, that are both optimized in terms of biological and therapeutic relevance as well as likelihood of success for developing peptide inhibitors. Mainly, we will integrate a number of different biological variables into an overall score. Among these variables is the output of aim 1 will serve as another feature in this algorithm. This aim will have two deliverables: First a list of potential target candidates as well as an optimized procedure to find such proteins.

Aim 3: Develop peptide inhibitors to the selected targets. Using modern phage display technology, we will obtain peptides that bind the selected targets with high affinity. Briefly, we will make use of the aforementioned computational prediction to generate optimized libraries for each target. Using these libraries, we will carry out several rounds of selection against the selected targets to generate high-affinity peptides. We can do these screens at a large scale, doing 96 of these screens at once. Then, we will validate the binding of these peptides in competitive binding experiments. Finally, we will confirm inhibition of biochemical activity of a selected set of the proteins with various collaborators.

Aim 4: Validate these inhibitors using a lentiviral delivery system. In the third stage of this project, we will test the biological efficacy of the developed peptides using a lentiviral delivery system. Briefly, we will clone selected peptides into a lentiviral vector, which we can then use to deliver it to the cancer cell line of choice. A reduced survival of selected cell lines will validate the peptides biological efficacy.

IV SCIENTIFIC SUMMARY

(Maximum one half (½) page)

Provide a brief summary of the research project. Include a description of (i) the proposed research, (ii) how this project is innovative, (iii) the expected outcomes, and (iv) future commercialization plans that would be based on this early-stage research. Please note if the project is successful in the competition, an additional Lay Summary will be required by project start for public distribution.

Protein interactions play a central role in many processes in the cell. Of particular importance for signaling pathways are interactions of modular protein domains. Aberrations in these domains, for instance in kinases, are very often a cause for cancer. Approaches based on combinatorial chemistry offer a promising venue to both detect and disrupt these interactions, leading to new insight into cancer biology.

Objectives: We propose to develop a novel phage-display based screening technology that instead of randomized libraries uses large well-defined libraries where each peptide has been designed.

Custom oligonucleotide microarrays are commercially available from a number of vendors. In preliminary experiments, we have established that the sequences from the microarrays are correctly incorporated into the phage genome and translated as peptides on the phage surface. We will computationally design the oligonucleotides to tile all disordered regions of the human proteome, such that potential binding sites to all possible peptide binding domains are represented in the library. This is in contrast to conventional experiments that either use randomized libraries or can only represent a small portion of the human proteome. In effect, our phage library will display all human binding motifs on its surface. We will then use this library to screen several hundred peptide binding domains to identify their binding partners in the human proteome. One application of this technology is to detect real interactions in human signaling pathways and uncover interaction motifs that remained hitherto undetected. In the future, we can extend it to screen other proteomes, e.g., of human pathogens. We plan to commercialize this technology as a service and offer the library for sale as a research reagent.

VI SCIENTIFIC SUMMARY (Maximum one (1) page)

Provide a scientific summary of the research proposed, which describes the overall goal(s) of the project, supporting objectives, anticipated outcomes, benefits for Canada and a summary of the integrated GE³LS research activities.

Human antibodies have emerged as major therapeutic agents for the treatment of cancer and other devastating diseases, and the market for antibody therapeutics currently stands at more than \$26B annually. Our major goal is to develop a program for efficient and large-scale production of synthetic antibody reagents. To harness the power of this program for the benefit of human health and the Ontario economy, we will apply the technology to develop antibodies against high-value cancer targets. Ultimately, the program will yield candidate therapeutic antibodies for commercialization. Indeed, the role of antibodies in diagnostics and as probes to understand cell biology and pathology will undoubtedly continue to expand, presenting another opportunity for commercialization.

Our goals will be achieved through collaborations between the Toronto Recombinant Antibody Technology Centre (TRAC) and cancer biologists associated with the Selective Therapies Program (STP) of the Ontario Institute for Cancer Research (OICR). We will create an advanced discovery pipeline for development and validation of novel synthetic antibody therapeutics, which includes:

Target identification: We have generated a panel of more than one hundred secreted proteins that are associated with cancer pathways and are within the research expertise of our cancer biology team. We will mine existing genomic and functional genomic data to further expand and refine our target list.

Antigen production: We will establish several methodologies for the rapid production of antigens for use in antibody selections and screens. Soluble ligands and receptor domains will be purified as recombinant proteins or will be displayed on yeast cells. Integral membrane proteins will be expressed in mammalian cells to enable antibody selections directly against cell lines.

Antibody generation: Using phage display technology, we will develop antibodies against antigens representing our panel of serum-accessible, cancer-associated proteins. We have established selection methodologies whereby naïve antibody-phage libraries can be used in selections against purified protein antigens or against antigens displayed on yeast or mammalian cell lines.

Antibody production: Antibodies will be purified directly from bacteria as monovalent antigen-binding fragments (Fabs). In addition, we have established a modular reformatting system that enables the rapid production of full-length, bivalent antibodies in the IgG format, the preferred format for therapeutics.

Antibody validation: Antibodies will be assayed for affinity and specificity with standardized in vitro assays. Antibodies that exhibit high affinity and specificity will be examined in cell-based assays to identify antibodies that (1) recognize native antigen in the cellular context, (2) inhibit downstream signalling events, and (3) inhibit the proliferation of cancer cell lines.

Evaluation of candidate therapeutic antibodies in xenografts: The most promising antibodies from the validation phase will be tested in mouse xenograft models to identify potential cancer therapeutics.

Our ultimate aim is to produce a panel of high value human antibody therapies suitable for clinical testing in patients. Our efforts will also produce highly valuable resources for drug discovery and validation with substantial commercial opportunities.

Philip M. Kim, PIN: 230541 – Form 101 – NSERC Discovery Grant – Proposal

NSERC Discovery Grant – Proposal: Analysis of Evolutionary Mechanisms in Signaling Networks

1. Summary, Background and Significance:

Recent years have seen a revolution in the study of biology. Novel technology in data acquisition has enabled a new systems level study of cells and organisms [1]. Maps of biomolecular networks emerged, which describe the complex interplay of genes, proteins and other biomolecules to together form a living being [2]. These networks will ultimately enable a fundamental understanding of biological systems from an engineer's perspective – the ability to accurately predict the effects of perturbations and the ability to design systems *do novo*. While there have been many great advances in the study of these complex biomolecular networks, very little is known about the principles that govern their evolution. How these highly complex and often tightly regulated systems evolved and which rules have shaped them remains an enigma to this day. The main reason for this is the absence of reliable comparative data. In this project, I will generate accurate and reliable network data on signaling networks for different organisms. This will allow me to rigorously analyze evolutionary processes in signaling networks and obtain fundamental insight into the evolution of higher organisms.

2. Objectives:

The overall goal of the proposed work is to find common rules that govern the evolution of cellular networks and in particular, of signaling networks. I propose to achieve this goal both through the development and application of a variety of novel computational techniques. This will allow me to, for the first time, perform a rigorous comparative study of interaction networks in different organisms.

2.1 Aim 1 – Generation of high quality networks in different organisms

The combination of genome-wide specificity maps that I have generated with my close collaborators [A1][A0a] with a novel computational method I recently developed [A0b] allows for the accurate generation of signaling networks in a variety of different species independent of sequence data. I will adapt these methods for both cases and predict accurate networks in the yeast and primate lineages.

2.2 Aim 2 – Analysis of signaling network evolution and detection of dynamic network evolutionary patterns

Based upon existing research, I will develop methods to analyze the specific conservation and evolutionary rewiring patterns of cellular networks. Of particular interest are the kinds of interaction modules that tend to be conserved (i.e., of likely higher importance to the cell) and what kinds of rewiring patterns tend to occur more frequently in evolution.

2.3 Aim 3 – Combined analysis of protein and gene regulatory network evolution

Building on results from Aim 2, I will integrate gene regulatory networks into the analysis and perform a global analysis of regulatory mechanisms in the cell. This

Philip M. Kim, PIN: 230541 – Form 101 – NSERC Discovery Grant – Proposal

will extend the evolutionary network formalism to the concept of combined evolutionary network motifs.

3. Recent Progress

3.1 Theoretical work on networks

Tremendous advances in high-throughput biology have enabled the mapping of various cellular networks, most importantly protein-protein interaction networks [3], gene regulatory networks [4], metabolic networks [5] and genetic interaction networks [A0c]. I have considerable experience in analyzing the topology of biological networks. For instance, with a collaborator, I introduced the novel concept of “bottlenecks”, i.e. nodes that are important not through their local connectivity (“hubs”), but through their centrality in the global network [A12]. In other recent work, I pioneered the use of information from protein 3D structures in the analysis of cellular networks [A13]. Extending this approach, I analyzed modern networks in terms of intrinsic disorder [A7]. Going further, I participated in a project to propose an ontology of edges to expand the analysis of networks [A11].

3.2 Signaling networks

Of particular importance for the cell are signaling networks, as they are central in the regulation of most, if not all, cellular processes. Recent high-throughput methods have started to map out some of these signaling networks [6], but many still rely on more traditional views of pathways and cascades. Signaling networks consist of protein modules that have undergone considerable expansion in recent evolutionary history [7]. Two particular important examples are protein kinases and peptide recognition domains (PRD). Both have in common that they bind short stretches of peptides in their target proteins with high specificity. These stretches are also known as linear motifs and play important roles in signaling themselves [8]. The mixing and matching of domain binding specificity with linear motifs in their targets is what determines the signaling network architecture and thereby their function. This particular feature makes them an attractive and tractable target for evolutionary study – we can map interactions in every organism with the genome sequenced, independent of homology assignment, by obtaining the binding specificity of all PRDs/kinases.

3.3 Initial studies on network evolution

There has been a considerable amount of work on theoretical studies of network evolution [9]. In the case of gene regulatory networks, comparative experimental studies have been carried out [10], identifying several mechanisms of evolution and enabling further theoretical work. However, for protein networks, studies have been hampered by the lack of comparative network data. More recently, specific aspects of network evolution have been analyzed, namely the evolution of protein kinases. Two recent studies have examined another specific sub-problem of network evolution, which is the evolution of phosphosites [11][12]. Their main finding was that phosphosites cluster and change quite rapidly. Many questions about the main mechanisms of pathway evolution remain, which I will address in this research project.

3.4 Novel experimental methods and data sets

INTRODUCTION

The current generation in the developed world faces tremendous sociological and economic challenges stemming from an ageing population. Perhaps the largest of these is the growing need for healthcare with, e.g., the incidence rates of many cancers rapidly increasing. To further compound this, the de novo development of drugs is slowing down due to the immense investments needed. Genomic science can help address these problems in two ways: First, by developing a greater understanding of genetic differences, predispositions to diseases and suitability to different drugs, we can build a predictive, preventative, personalized medicine, thereby leading to a much more effective healthcare system.

Second, by modeling the complex pathways in humans and mapping the genetic variants and mutations involved in different diseases and cancers, we can generate new hypotheses about drug targets and, in collaboration with clinical researchers, greatly accelerate the development of new drugs.

GENERAL DESCRIPTION OF INFRASTRUCTURE

The requested infrastructure consists of high-performance servers, workstations, laptop computers and software. This infrastructure will be the main workhorse for Dr. Kim's group's activities in research and training. Furthermore, it will support many research groups in and outside of Toronto indirectly, by providing Dr. Kim with the ability to collaborate with them. Specifically, the infrastructure consists of a total of seven Xserve high-performance machine servers, two 16TB RAID storage arrays, eight MacPro workstations and one MacBook Pro laptop computer. In addition, software and support equipment as well as support maintenance costs are needed.

LOCATION OF INFRASTRUCTURE

The Kim lab is housed inside the Terrence Donnelly Centre for Cellular and Biomolecular Research (CCBR) at the University of Toronto. The sixth floor of the CCBR is the Centre for Computational Biology and Dr. Kim's office and his lab are located there. The workstations will be located in the 6th floor 20,000ft. open concept laboratory space, whereas the servers and the storage array will be located in a dedicated server room, previously outfitted using funding from the CFI and Genome Canada grants.

RESEARCH ENABLED BY INFRASTRUCTURE

The current rollout of next generation sequencing and high-throughput genotyping technologies is leading to a revolution in the area of genetics. This, in turn, will have many effects on other fields of biomedical research. In particular, it will enable the long-envisioned personalized, genomic medicine. Much progress has already been made in mapping genetic variants. In fact, completely new phenomena of human genetic variation have been discovered in the past few years. Moreover, genetic variation has been mapped extensively to phenotypic features, such as disease susceptibility. This trend is going to continue and, in fact, greatly accelerate in the next few years. These approaches work very well in

mapping variation, but do not provide much mechanistic insight, and hence medical applicability in a predictive context. This is despite a substantial body of accumulated knowledge in various disciplines such as genomics, systems biology and structural biology. By using modern Machine Learning tools and leveraging interdisciplinary expertise, we can build tools that make use of this knowledge. Specifically, we can integrate information from protein networks and protein structures, as well as other genomic features. By expanding the approach beyond the one gene paradigm we can greatly aid genotype-phenotype predictions and gain mechanistic understanding of phenotype causing effects.

IMPORTANCE OF THE RESEARCH TO CANADA

Short-term benefits include, but are not limited to, new advances in Canadian biomedical science, more efficient use of research resources in Canada, and more training of highly qualified personnel. In the longer term, the proposed research will make significant advances towards personalized and predictive medicine, thereby greatly aiding the Canadian health care system and also promoting Canada's position in this field of future economic growth.

CURRENT NEED FOR INFRASTRUCTURE

Dr. Kim started as a new faculty member at the CCBR and the Banting and Best Department of Medical Research in January 2009. The requested infrastructure will serve as the backbone of computing equipment in his laboratory.

CREATING TRAINING OPPORTUNITIES

Training of highly qualified personnel is integral to this proposal. The funding requested here will increase the number of people Dr. Kim will be able to train more than threefold, from two to eight, his goal is to train five doctoral students, one postdoctoral researcher and two undergraduate research assistants.

STRENGTHENING EXISTING AND PLANNED COLLABORATIONS

The requested infrastructure will greatly help current and planned collaborations by providing Dr. Kim's group with the computing power to carry out the extensive calculations needed by his collaborators who are carrying out studies on genetic variants and protein interaction networks.

USE OF EXISTING INFRASTRUCTURE

The servers will be located in a shared server room that was previously built using funding from Genome Canada (PI: Dr. Gary Bader and Dr. Quaid Morris). Excess capacity on two large computing clusters that were purchased with Genome Canada funding will be used for large production-style computation tasks. Furthermore, Gigabit Ethernet infrastructure previously installed by funding also from Genome Canada will be utilized for stable and fast network connectivity.