



Centre for
Comparative **G**enomics &
Evolutionary **B**ioinformatics
www.cgeb.dal.ca Dalhousie University, Halifax, NS, CANADA

ANNUAL REPORT

October 1, 2009 – October 1, 2010

Table of Contents

I.	Executive Summary	1-3
II.	Financial Status of the Centre	4-5
III.	CGEB Membership	6
IV.	CGEB Governance	7
V.	Supervision of Research Trainees	8-9
VI.	Publications by CGEB Faculty and Trainees	10-15
VII.	Presentations by CGEB Members	16-18
VIII.	Other Research Contributions from CGEB Trainees	19-23
IX.	Collaborative Interactions between CGEB Members	24-25
X.	CGEB Developed Software	26-29
XI.	Current External Research Grants and Funding	30-32
XII.	Honours/Awards/Distinctions	33
XIII.	Service to the Scientific Community (outside Dalhousie).	34-35
XIV.	Other Activities of the Centre	
	A. CGEB Seminar Series	36
	B. CGEB Meeting Sponsorships	36
	C. CGEB Joint Lab Meetings	37
	D. CGEB and SEB Journal Club.	37
XV.	Future Directions	38-39
	<i>APPENDIX A: CGEB Governance</i>	40-43

I. EXECUTIVE SUMMARY

Director: Dr. Andrew J. Roger
Phone: 494-2620
Fax: 494-1355
Coordinator: Wanda Danilchuk (494-3569)
Website: <http://www.cgeb.dal.ca>

The vast majority of biodiversity on Earth is microscopic and unicellular. These organisms play key roles in all known ecosystems and have done so since life first arose on the planet ~3.8 billion years ago. Microbes are the major primary producers in the ocean, they produce most of the oxygen we breathe, they are central players in nutrient cycling in terrestrial and aquatic habitats and are abundant commensals, symbionts and pathogens capable of causing disease in humans and other organisms. Yet despite their importance, until recently, very little was known about the ‘true’ diversity of unicellular organisms and their environmental impact beyond the few that were culturable in the laboratory. Research on microbial diversity and genomics in the last decade has revolutionized our understanding of the microbial world. It is now possible to isolate diverse unicellular organisms from the environment and characterize their genomes quickly and cheaply. Furthermore, meta-genomic technologies allow genomic surveys of whole microbial communities without the need for culturing at all. With the resulting exponential growth of genomic sequence databases and associated meta-data, development of computational and statistical methods for extraction of biologically meaningful information and hypothesis testing is becoming increasingly important. It is at this nexus of genomic, computational biological and evolutionary research that the internationally recognized researchers at the Centre for Comparative Genomics & Evolutionary Bioinformatics (CGEB) at Dalhousie University are making important advances.

The core of CGEB encompasses ten researchers in the Faculty of Medicine [Andrew Roger, Ford Doolittle, Michael Gray, John Archibald, Claudio Slamovits], Faculty of Science [Joseph Bielawski, Alastair Simpson, Ed Susko] and Faculty of Computer Science [Robert Beiko, Christian Blouin]. In the last year, CGEB has expanded to include three affiliated ‘CGEB Associate’ investigators from the Faculty of Science: two from Mathematics and Statistics [Hong Gu and Chris Field] and one from Biology [Bob Lee]. The main focus of the Centre’s research is understanding how microbial genomes evolve and diversify, however the research programs collectively span computational biology, computer science, statistical modeling and comparative genomics, with a strong focus on method and theory.

The Centre was officially approved by the Dalhousie Senate in June, 2008 with the mission to foster collaborations amongst a ‘group’ of nine comparative genomics/ bioinformatics and microbial evolution researchers at Dalhousie. Specifically, the CGEB Centre’s goals are to:

- 1) recruit new world-class comparative genomics/bioinformatics researchers to Dalhousie;
- 2) recruit top-notch postdoctoral fellows and Ph.D. students to CGEB labs through the provision of Tula Foundation stipends and research allowances;
- 3) create a CGEB seminar series to attract world-class speakers to interact with CGEB principal investigators and trainees;
- 4) fund CGEB trainees to attend and present their work at national and international scientific meetings; and
- 5) provide a framework that promotes collaborations between CGEB members and attract external research funding. *For each of the last two years, we have met or exceeded all of these goals.*

This Annual Report records the achievements of the Centre over the past year, the highlights of which are summarized below.

CGEB research productivity, funding and collaborations.

During the period October 2009-October 2010, CGEB researchers have published a total of 75 manuscripts in peer-reviewed journals (pp. 10-15), including contributions in high impact journals such as *Current Biology*, *Genome Research*, *Proceedings of the National Academy of Sciences USA*, *Science* and *Trends in Ecology and Evolution*. The international profile of CGEB researchers is further demonstrated by the 30 invited presentations given at international conferences or invited seminars over the past year (pp. 16-18). Furthermore, in the past year CGEB members have been awarded a number of prestigious Honours and Distinctions (page 33). The CGEB trainees (postdocs, graduate students and undergraduates) have also been productive beyond their contribution to publications (pp. 19-23). They have presented their results at international meetings (a total of 45 seminars and posters), and, in the process, gained invaluable presentation and networking skills. The CGEB Trainee Travel to Meetings funds provided by the Tula Foundation have been of great benefit in supporting the latter activities.

A large portion of CGEB research efforts is directed at development of computer software tools for evolutionary modeling analyses and bioinformatics. To date, CGEB has produced 21 bioinformatics software tools, six of these in the last year. These ‘open-source’ software tools (or suites of tools) for analysis of genomic data have been created, published and made available to the international scientific community through free web access (pp. 26-29).

CGEB researchers have also been extremely successful in attracting external funding, garnering a total of ~\$1,973,208 of external grant support in total over 2009-2010 from Natural Sciences and Engineering Research Council of Canada, Canadian Institutes for Health Research, Canada Foundation for Innovation, Canada Research Chairs (CIHR), the Canadian Institute for Advanced Research, as well as the Tula Foundation (pp. 30-32).

Finally, CGEB has greatly facilitated collaborative interactions between the member researchers leading to joint publications, software development and grant applications (pp. 24-25). The CGEB journal club, joint lab meetings, seminar series and, in future, the annual mixer/poster session will continue to foster CGEB collaborations and synergy.

CGEB postdoc and student recruitment and training.

Since we first received the trainee funds from the Tula Foundation we have recruited eight top-notch postdoctoral fellows and five Ph.D. students to fill these positions. The importance of these Tula-sponsored trainees to the success of our research programs cannot be over-stated. As these individuals are often already accomplished young scientists, they bring new techniques/knowledge to CGEB labs, as well as help train the younger, less experienced students in the labs (thereby gaining invaluable supervisory experience). This is especially important to the success of the CGEB labs because the total number of CGEB trainees collectively for 2009-2010 was 56 (see pp. 8-9).

CGEB seminar series.

The CGEB Seminar Series is in full swing and has hosted seven well-attended seminars this year from internationally recognized researchers (see p. 36) with several more scheduled to occur over the next few months. The seminar series has expanded to include sponsorship of symposia in meetings held at Dalhousie University by paying for the travel and accommodation of the

guest speakers. In return, CGEB members are free to choose the topic of the symposium and the speakers that will present. In June 2010 we sponsored symposium speakers at the CIFAR-IMB Workshop on Eukaryotic Genome Annotation (see p. 36).

Ongoing CGEB activities and future directions.

So far, group CGEB activities have consisted of two bi-weekly journal clubs, monthly joint lab meetings where CGEB PIs or trainees present their ongoing research, and the CGEB seminar series. However, we have decided that a more intense forum for interaction will be necessary to really generate and flesh out ideas for new large-scale CGEB-based collaborative programs. Thus, we have decided to hold an annual (or biennial) offsite 2-3 day ‘CGEB retreat’ with all CGEB PIs, Associates and trainees attending. Procuring the funds to support this retreat represents a current challenge. Another initiative is to sponsor an international meeting in comparative genomics, microbial biodiversity and phylogenomics to allow CGEB researchers and trainees to interact with other internationally renowned scientists. Finally, we have determined that building, maintaining and staffing a stable and sophisticated microbial genomics database, software repository and comparative genome analysis platform is common goal that would greatly benefit the CGEB collective and allow us to continue to be world-leaders in these fields.

These plans are discussed in more detail in the ‘Future Directions’ section (see pp. 38-39).

II. FINANCIAL STATUS OF THE CENTRE

The Centre's administration (i.e. 50% administrator's salary, seminar series fund, and student travel funds) is supported by funding from the Tula Foundation at \$50,000/year for 8 years, 2007-2015 (*see Table A. below*). The other 50% of the administrator's salary is covered by contributions from the Faculties of Medicine, Science and Computer Science, and the office of VP, Academic & Provost. Five-year financial commitments from Dalhousie were established and annual installments for years 1-3 have been received (*see Table B. below*).

A. Tula Foundation Contributions

Fiscal Year (August 1-July 30)	50% salary of administrator	Seminar Series Fund	Trainee Travel/ Meetings Fund	TOTAL
Year 1: 2007-2008	\$25,000	\$15,000	\$10,000	\$50,000
Year 2: 2008-2009	\$25,000	\$15,000	\$10,000	\$50,000
Year 3: 2009-2010	\$25,000	\$15,000	\$10,000	\$50,000
Year 4: 2010-2011	\$25,000	\$15,000	\$10,000	\$50,000
Year 5: 2011-2012	\$25,000	\$15,000	\$10,000	\$50,000
Year 6: 2012-2013	\$25,000	\$15,000	\$10,000	\$50,000
Year 7: 2013-2014	\$25,000	\$15,000	\$10,000	\$50,000
Year 8: 2014-2015	\$25,000	\$15,000	\$10,000	\$50,000
Total:				\$400,000

The *Seminar Series Fund* enables us to invite world-renowned scientists to present seminars at Dalhousie University. The *Travel to Meetings Fund* enables our CGEB trainees to attend scientific meetings and workshops.

The Centre is also supported by fellowships from the Tula Foundation that are eligible for each CGEB faculty member for postdoctoral trainees (\$44,000/year) or graduate student trainees (\$25,000/year). In addition, the fellowships also provide a research allowance for each trainees' lab (\$20,000/year for molecular biology PDFs, \$5,000/year for computational PDFs, \$2,500/year for computational PhDs). Each fellowship is awarded for five years.

In summary, the Tula Foundation has invested over \$3,000,000 in the CGEB centre.

B. Dalhousie Faculty Contributions

Fiscal Year (April 1-March 31)	Faculty of Medicine	VP, Academic & Provost	Faculty of Science	Faculty of Comp. Sci.	TOTAL
Year 1: 2008-2009	\$15,000	\$10,000	N/A	N/A	\$25,000 (received)
Year 2: 2009-2010	\$15,000	\$10,000	\$1,000	\$1,000	\$27,000 (received)
Year 3: 2010-2011	\$10,000	\$10,000	\$6,000	\$3,000	\$29,000 (received)
Year 4: 2011-2012	\$10,000	\$10,000	\$7,000	\$4,000	\$31,000
Year 5: 2012-2013	\$10,000	\$10,000	\$8,000	\$5,000	\$33,000
Total:					\$145,000

C. CGEB New Opportunities Fund

Another source of funds available to finance CGEB activities is our *New Opportunities Fund*. This is a “special purpose account” (# 68292) that was opened to receive transfers of revenue that was generated from CGEB faculty activities. The purpose of this fund is to provide additional partial funding or ‘seed’ monies for CGEB activities, the costs of which either exceed existing allocated funds (e.g. the Trainee Travel/Meeting fund, the Seminar Series fund) or one-time events/activities/opportunities that fall under the CGEB mandate. Through this fund CGEB co-sponsored the *CIFAR-IMB Workshop on Eukaryotic Genome Annotation* at Dalhousie in 2010 with the Canadian Institute for Advanced Research (CIFAR) (see p. 36 for details under “CGEB Meeting Sponsorship”). To date these funds have also been used for costs associated with recruitment of potential trainees to CGEB labs (e.g. interview travel and accommodation costs), as well as the initial CGEB website design (see <http://www.cgeb.dal.ca>).

The first source of revenue for this fund came from the *Society for Molecular Biology and Evolution Conference* that resulted in a net income of \$39,181.46. Second, Drs. Doolittle and Gray had been maintaining an account from an earlier conference that had generated a net revenue of \$10,073.15. In 2008 they generously donated these funds to the CGEB New Opportunities Fund. Finally, another symposium (“*Mitochondria, Ribosomes & Cells*” in 2008) generated a net revenue of \$106.01.

The current balance in our CGEB New Opportunities Fund is **\$40,123.29**.

In keeping with the wishes of the original donors, we have been developing plans to use these monies to help leverage funds to promote more interaction/collaborations within CGEB and between CGEB members and researchers worldwide. Some of these plans are detailed in the “Future Directions” section (see pp. 38-39).

III. CGEB MEMBERSHIP

DIRECTOR:

- Andrew Roger, Professor, Biochemistry and Molecular Biology

STEERING COMMITTEE:

- Andrew Roger, Professor, Biochemistry and Molecular Biology
- Michael W. Gray, Professor Emeritus, Biochemistry and Molecular Biology
- W. Ford Doolittle, Professor Emeritus, Biochemistry and Molecular Biology

CGEB FACULTY MEMBERS:

- John Archibald, Associate Professor, Biochemistry and Molecular Biology
- Robert Beiko, Assistant Professor, Computer Science
- Christian Blouin, Associate Professor, Computer Science, and Biochemistry and Molecular Biology
- Joseph Bielawski, Associate Professor, Biology, and Mathematics and Statistics
- Ford Doolittle, Professor Emeritus, Biochemistry and Molecular Biology
- Michael Gray, Professor Emeritus, Biochemistry and Molecular Biology
- Andrew Roger, Professor Biochemistry and Molecular Biology
- Alastair Simpson, Associate Professor, Biology
- Claudio Slamovits, Assistant Professor, Biochemistry and Molecular Biology
- Ed Susko, Professor, Mathematics and Statistics

ASSOCIATES:

- Christopher Field, Professor Emeritus, Mathematics and Statistics
- Hong Gu, Associate Professor, Mathematics and Statistics
- Robert Lee, Adjunct Professor, Biology

CENTRE ADMINISTRATOR:

- Wanda Danilchuk, Biochemistry and Molecular Biology

IV. CGEB GOVERNANCE

CGEB Governance structure and procedures have been working smoothly since the creation of the Centre and the update of procedures last year (*see Appendix A*). As the Centre has only ten current members plus three associates and maintains no core Centre-specific research facilities or equipment, governance is not onerous and no problems with our procedures have surfaced to date. According to our governance guidelines, the CGEB Director is to meet annually with the ‘Deans Committee’ for discussion regarding Centre operations and directions. The Deans’ committee meeting will be arranged to occur in November, 2010. Over the past year, we have had regular CGEB members meetings and have discussed issues related to Centre structure and activities. These are briefly summarized below:

A. CGEB Associates

Three associates were appointed in 2010:

- Dr. Hong Gu (Associate Professor, Mathematics and Statistics)
- Dr. Chris Field (Professor Emeritus, Mathematics and Statistics)
- Dr. Robert Lee (Adjunct Professor, Biology)

B. Opportunities and challenges to be overcome in the future development of CGEB

A common research theme of CGEB group is the gathering, annotation, visualization, analysis and archiving of enormous quantities of microbial genome sequence data. Because of the rapid growth in the development of high-throughput next generation sequencing technology and its ever-lowering costs, we are entering a phase where analysis of the data has become **the** limiting step in the scientific process. While we have been successful in procuring grant funds to generate genomic data both within and outside Canada, it is becoming increasingly difficult to secure monies for the trainees, staff and computational infrastructure that support the critical analysis of these data. Indeed several major database efforts in Canada that were uniquely serving these purposes (including TBestDB and GOBASE at Université de Montréal) have recently ‘died’ as a result of the lack of infrastructure funding. Although CGEB groups have individually cobbled together funds to support smaller ‘in-house’ database/analysis pipeline systems based on individual trainee efforts, it would be much more productive to take advantage of our unique strengths and expertise in microbial genomics by developing a central web-accessible CGEB comparative genomics database and analysis ‘platform’. Indeed, Beiko has been developing his MOA system in this direction as collaborations between Beiko’s group and other CGEB labs have become established (e.g. Archibald and Gray). The barrier to this effort however remains our ability to fund dedicated staff members whose job is to expand and maintain the MOA database and populate it with data from the CGEB labs, support the many software tools we have developed and administer the huge computational resources upon which the MOA system and these other software tools depend (e.g. the 3 large computer clusters within CGEB). Until now the development and maintenance of these resources have depended on CGEB PIs, students and postdocs whose time is limited and would be much better spent doing research. The lack of funding for dedicated ‘cyber-infrastructure’ staff and the dim prospects for obtaining it through granting agencies (who prefer to fund trainees) is the biggest challenge for the future growth of CGEB research efforts in comparative genomics.

V. SUPERVISION OF RESEARCH TRAINEES

A total of **56** trainees were supervised in CGEB members' labs (10/2009 – 10/2010).

Note: CGEB trainees that are specifically **Tula-funded** are highlighted in bold

CGEB Faculty Member:	Research Trainee:
Andrew Roger Biochemistry and Molecular Biology	<p>Matt Brown, Postdoctoral Fellow Eleni Gentekaki, Postdoctoral Fellow Huaichun Wang, Postdoctoral Fellow (<i>co-supervised with E. Susko</i>) Anastasios Tsaousis, Postdoctoral Fellow Tommy Harding, PhD student Michelle Leger, PhD student Dan Gaston, PhD student Martin Kolisko, PhD student (<i>co-supervised with A. Simpson</i>) Liwen Zou, PhD student (<i>co-supervised with E. Susko</i>) Courtney Stairs, PhD student Javier Alfaro, MSc student Grant Stevens, Honours and summer student Jitka Hlavackova, Visiting PhD student (Czech Republic) Kiyotaka Takishita, Visiting Researcher (JAMSTEC, Japan)</p>
John Archibald Biochemistry and Molecular Biology	<p>Goro Tanifuji, Postdoctoral Fellow (<i>co-supervised with M.W. Gray</i>) Eunsoo Kim, Postdoctoral Fellow Julia Hopkins, Postdoctoral Fellow Shinichiro Maruyama, Visiting Researcher (JSPS, Japan) Bruce Curtis, PhD student Robert Eveleigh, MSc student (<i>co-supervised with R. Beiko</i>) Christa Moore, MSc student Tyler Mills, Honours student Rebecca Gibault, Honours student Katherine Richman, Honours student</p>
Joseph Bielawski Biology, and Mathematics and Statistics	<p>Katherine Dunn, Postdoctoral Research Associate Rachael Bay, MSc student Caroline Urqhardt, MSc student Joseph Mingrone, MSc student (<i>co-supervised with E. Susko</i>) Yuriy Khalak, Undergraduate student</p>
Christian Blouin Computer Science, and Biochemistry and Molecular Biology	<p>Jose Sergio Hleap Lozano, PhD student Swati Shankar, PhD student Haibin Liu, PhD student</p>

CGEB Faculty Member	Research Trainee
Robert Beiko, Computer Science	Dennis Wong, PhD student
	Norman MacDonald, PhD student
	Christopher Whidden, PhD student
	Donovan Parks, PhD student
	Robert Eveleigh, Master's student (<i>co-supervised with J. Archibald</i>)
	Scott Perry, Master's student
	Sylvia Churcher, Master's student (<i>co-supervised with C. Field</i>)
	Christopher Smith, Master's student
	Mike Porter, Master's student
	Catherine Holloway, NSERC USRA student
W. Ford Doolittle Biochemistry and Molecular Biology	Olga Zhaxybayeva, Postdoctoral Research Associate
Michael Gray Biochemistry and Molecular Biology	Goro Tanifuji, Postdoctoral Fellow (<i>co-supervised with J. Archibald</i>)
	Ryan Gawryluk, PhD student
Alastair Simpson, Biology	Jong Soo Park, Postdoctoral Fellow
	Aaron Heiss, PhD student
	Martin Kolisko, PhD student (<i>co-supervised with A. Roger</i>)
	Tommy Harding, PhD student (<i>co-supervised with A. Roger</i>)
	Tyler Wilding, Honours student
	Karla Grimm, Honours student
	Greg Macmullin, Undergraduate summer student
Qianqian Zhang, Visiting Scholar (VSGS student)	
Claudio Slamovits Biochemistry and Molecular Biology	Shehre-Banoo Malik, CIFAR Junior Fellow
	Gillian Gile, Postdoctoral Fellow
	Renny Lee, PhD student
	Hugo Lai, Honours student
Edward Susko Mathematics and Statistics	Huaichun Wang, Postdoctoral Fellow (<i>co-supervised with A. Roger</i>)
	Liwen Zou, PhD student (<i>co-supervised with A. Roger</i>)
	Jihua Wu, PhD student
	He Gao, MSc student
	Joey Mingrone, MSc student (<i>co-supervised with J. Bielawski</i>)
	Goldis Radajapalipour, MSc student

VI. PUBLICATIONS BY CGEB FACULTY AND TRAINEES

(includes published, in press & accepted papers: total of 75 for Oct. 2009-Oct. 2010)

Note: for each publication, CGEB faculty members are highlighted in bold and CGEB trainees' names are underlined.

Bae, Y., Kim, E., L'Hernault, S.W. and Barr, M.M. (2009) The CIL-1 PI 5-phosphatase localizes TRP polycystins to cilia and activates sperm in *C. elegans*. *Current Biology* 19: 1599–1607.

Bapteste, E., O'Malley, M.A., **Beiko, R.G.**, Ereshefsky, M., Gogarten, J.P., Franklin-Hall, L., Lapointe, F.J., Dupré, J., Dagan, T., Boucher, Y. and Martin W. (2009) Prokaryotic evolution and the tree of life are two different things. *Biology Direct* 4: 34-54.

Baurain, D., Brinkmann, H., Petersen, J., Rodríguez-Ezpeleta, N., Stechmann, A., Demoulin, V., **Roger, A.J.**, Burger, G., Lang, B.F. and Philippe, H. (2010) Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes and stramenopiles. *Molecular Biology and Evolution* 27: 1698-1709.

Beiko, R.G. (2010) Gene sharing and genome evolution: Networks in trees and trees in networks. *Biology and Philosophy* 25: 659-673.

Blouin, C., Perry S., Lavell A., **Susko E.** and **Roger A.J.** (2009) Reproducing the manual annotation of multiple sequence alignment using a SVM classifier. *Bioinformatics* 25: 3093-3098.

Bullerwell, C.E., Burger, G., Gott, J.M., Kourennaia, O., Schnare, M.N. and **Gray, M.W.** (2010) Abundant 5S rRNA-like transcripts encoded by the mitochondrial genome in amoebozoia. *Eukaryotic Cell* 9: 762-773.

Carlton, J.M., Malik S.-B., Sullivan S.A., Tang P. and R.P. Hirt (2010) Chapter 3: The genome of *Trichomonas vaginalis*. In *Anaerobic Parasitic Protozoa: Genomics and Molecular Biology*, C.G. Clark, P.J. Johnson & R.D. Adam (Eds.), Caister Academic Press.

Chan, C.X., **Beiko, R.G.**, Darling, A.E. and Ragan, M.A. (2009) Lateral transfer of genes and gene fragments in prokaryotes. *Genome Biology and Evolution* 1: 429-438.

Charette, J.M. and **Gray, M.W.** (2009) U3 snoRNA genes are multi-copy and frequently linked to U5 snRNA genes in *Euglena gracilis*. *BMC Genomics* 10: 528.

Curtis, B.A. and **Archibald, J.M.** (2010) A spliceosomal intron of mitochondrial DNA origin. *Current Biology*, in press.

Curtis, B.A. & **Archibald, J. M.** (2010) Problems and progress in understanding the origins of mitochondria and plastids. In: *Symbiosis and Stress*, J. Seckbach & M. Grube (Eds.), Springer-Verlag, in press.

- Donaher, N., Tanifuji, G., Onodera, N.T., Malfatti, S.A., Chain, P.S.G., Hara, Y. and **Archibald, J.M.** (2009) The complete plastid genome sequence of the secondarily non-photosynthetic alga *Cryptomonas paramecium*: reduction, compaction, and accelerated evolutionary rate. *Genome Biology and Evolution* 1: 439-448.
- Doolittle, W.F.** (2010) The attempt on the life of the Tree of Life: science, philosophy and politics. *Biology and Philosophy* 25: 455-473.
- Doolittle, W.F.** and Zhaxybayeva, O. (2010) Metagenomics and the units of biological organization. *Bioscience* 60: 102-112.
- Dunn, K.A., **Bielawski, J.P.**, Ward, T.J., Urquhart, C. and **Gu, H.** (2009) Reconciling ecological and genomic divergence among lineages of listeria under an “extended mosaic genome concept”. *Molecular Biology and Evolution* 26: 2605-1615.
- Elias, M. and **Archibald, J.M.** (2009) Sizing up the genomic footprint of endosymbiosis. *Bioessays* 31: 1273-1279.
- Gawryluk, R.M. and **Gray, M.W.** (2010) An ancient fission of mitochondrial cox1. *Molecular Biology and Evolution* 27: 7-10.
- Gawryluk, R.M. and **Gray, M.W.** (2010) Evidence for an early evolutionary emergence of gamma-type carbonic anhydrases as components of mitochondrial respiratory complex I. *BMC Evolutionary Biology* 10: 176.
- Graham, L.E., Kim, E., Arancibia-Avila, P., Graham, J.M. and Wilcox, L.E. (2010) Evolutionary and ecophysiological significance of sugar utilization by the peatmoss *Sphagnum compactum* (Sphagnaceae) and the common charophycean associates *Cylindrocystis brebissonii* and *Mougeotia* sp. (Zygnemataceae). *American Journal of Botany* 97: 1–7.
- Gott, J.M., Somerlot, B.H. and **Gray, M.W.** (2010) Two forms of RNA editing are required for tRNA maturation in *Physarum* mitochondria. *RNA* 16: 482-488.
- Gray, M.W.** (2010) Re-thinking plastid evolution. *EMBO Reports* 11: 562-563.
- Gray, M.W.** (2010) Organelles. In *Brenner's Online Encyclopedia of Genetics 2E*, S. Maloy and K. Hughes (eds.), Academic Press, *in press*.
- Gray, M.W.** (2010) Mitochondria. In *Brenner's Online Encyclopedia of Genetics 2E*, S. Maloy and K. Hughes (eds.), Academic Press, *in press*.
- Gray, M.W.** (2010) Mitochondrial DNA. In *Brenner's Online Encyclopedia of Genetics 2E*, S. Maloy and K. Hughes (eds.), Academic Press, *in press*.
- Gray, M.W.** (2010) Mitochondrial genome. In *Brenner's Online Encyclopedia of Genetics 2E*, S. Maloy and K. Hughes (eds.), Academic Press, *in press*.

- Gray, M.W.** (2010) Mitochondrial mutants. In *Brenner's Online Encyclopedia of Genetics 2E*, S. Maloy and K. Hughes (Eds.), Academic Press, *in press*.
- Gray, M.W.** and Beyer, A.L (2010) Ribonucleic acid (RNA). In *McGraw Hill Encyclopedia of Science & Technology*, *in press*.
- Gray, M.W.** (2009) The path to RNA editing in plant mitochondria: the Halifax chapter. *IUBMB Life* 61: 1114-1117.
- Gray, M.W.** (2009) RNA editing in plant mitochondria: twenty years later. *IUBMB Life* 61: 1101-1104.
- Gu, H., Dunn, K.A. and **Bielawski, J.P.** (2010) Likelihood based clustering (LiBaC) for codon models, In: *Codon Evolution*, G.M. Cannarozzi and A. Schneider (Eds.), Oxford University Press, *in press*.
- Heiss, A.A., Walker, G. and **Simpson, A.G.B.** (2010) Clarifying the taxonomic identity of a phylogenetically important group of eukaryotes: *Planomonas* is a junior synonym of *Ancyromonas*. *Journal of Eukaryotic Microbiology* 57: 285-293.
- Heiss, A.A., Walker, G. and **Simpson, A.G.B.** (2010) The ultrastructure of *Ancyromonas*, a eukaryote without supergroup affinities. *Protist*, *in press*.
- Hjort, K., Goldberg, A.V., Tsaousis, A.D., Hirt, R.P. and Embley, T.M. (2010). Diversity and reductive evolution of mitochondria among microbial eukaryotes. *Philosophical Transactions of the Royal Society London B: Biol Sci* 365: 713-727.
- Hoef-Emden, K. and **Archibald, J.M.** (2010) Phylum Cryptophyta (Cryptomonads). In: *Handbook of Protozoology*, 2nd Ed., Margulis, L.(ed.), Jones & Bartlett Publishers, *in press*.
- Hopkins, J. and **Archibald, J.M.** (2010) Plastid evolution and the nuclear genomic 'footprint' of red and green algal endosymbionts. In: *Red Algae in the Genomics Age*, J. Seckbach & M. Grube, M. (Eds.), Springer-Verlag, pp. 191-204.
- Hug, L.A., Stechmann, A. and **Roger, A.J.** (2010) Phylogenetic distributions and histories of proteins involved in anaerobic pyruvate metabolism in eukaryotes. *Molecular Biology and Evolution* 27: 311-324.
- Kamikawa, R., Sanchez-Perez, G.F., Sako, Y., **Roger, A.J.** and Inagaki, Y. (2009) Expanded phylogenies of canonical and non-canonical types of methionine adenosyltransferase reveal a complex history of these gene families in eukaryotes. *Molecular Phylogenetics and Evolution* 53: 565-570.
- Kim, E. and **Archibald, J.M.** (2010) Plastid evolution: gene transfer and the maintenance of 'stolen' organelles. *BMC Biology* 8: 73.
- Kim, E., Park, J.S., **Simpson, A.G.B.**, Matsunaga, S., Watanabe, M., Murakami, A., Sommerfeld, K., Onodera, N.T. and **Archibald, J.M.** (2010) Complex array of endobionts

- in *Petalomonas sphagnophila*, a large heterotrophic euglenid protist from Sphagnum-dominated peatlands. *ISME Journal* 4: 1108-1120.
- Kim, E., Yubuki, N., Leander, B.S. and Graham, L.E. (2010) Ultrastructure and 18S rDNA phylogeny of *Apoikia lindahlii* comb. nov. (Chrysophyceae) and its epibiotic protists, *Filos agilis* gen. et sp. nov. (Bicosoecida) and *Nanos amicus* gen. et sp. nov. (Bicosoecida). *Protist* 161: 177–196.
- Kolisko, M., Silberman, J.D., Cepicka, I., Yubuki, N., Takishita, K., Yabuki, A., Leander, B.S., Inouye, I., Inagaki, Y., **Roger, A.J.** and **Simpson, A.G.B.** (2010) A wide diversity of previously undetected free-living relatives of diplomonads isolated from marine/saline habitats. *Environmental Microbiology* [May 13 - Epub ahead of print, PMID 20482740].
- Liu, H., **Blouin, C.** and Keselj, V. (2010) Identifying Interaction Sentences from Biological Literature Using Automatically Extracted Patterns. *Data & Knowledge Engineering, in press*.
- Lowe, C.D., Keeling, P.J., Martin, L.E., **Slamovits, C.H.**, Watts, P.C. and Montagnes, D.J.S. (2010) Who is *Oxyrrhis marina*? Morphological and phylogenetic studies on an unusual dinoflagellate. *Journal of Plankton Research, accepted for publication*.
- Lynn, D.H. and **Simpson, A.G.B.** (2010) From the editors: Describing new taxa of unicellular protists. *Journal of Eukaryotic Microbiology* 56: 403-405.
- MacDonald, N.J. and **Beiko, R.G.** (2010) Efficient learning of microbial genotype-phenotype association rules. *Bioinformatics* 26: 1834-1840.
- Martin, D.P. and **Beiko, R.G.** (2010) Genetic recombination and bacterial population structure. In: *Bacterial Population Genetics in Infectious Disease*, D.A. Robinson, D. Falush, and E.J. Feil (Eds.), Wiley, *in press*.
- Moore, C.E. and **Archibald, J.M.** (2009) Nucleomorph genomes. *Annual Review of Genetics* 43: 251-264.
- Nesbø, C.L., Kumaraswamy, R., Dlutek, M., **Doolittle, W.F.** and Foght, J. (2010) Searching for mesophilic Thermotogales bacteria: "mesotogas" in the wild. *Applied and Environmental Microbiology* 76: 4896-4900.
- Park, J.S. and **Simpson, A.G.B.** (2010) Characterization of halotolerant Bicosoecida and Placididea (Stramenopila) that are distinct from marine forms, and the phylogenetic pattern of salinity preference in heterotrophic stramenopiles. *Environmental Microbiology* 12:1173-1184.
- Park, J.S., Kolisko, M. and **Simpson, A.G.B.** (2010). Cell morphology and formal description of *Ergobibamus cyprinoides* n. gen., n. sp., another *Carpediemonas*-like relative of diplomonads. *Journal of Eukaryotic Microbiology (in press)*.

- Park, J.S., Vreeland, R.H., Cho, B.C., Lowenstein, T.K., Timofeeff, M.N. and Rosenzweig, W.D. (2009) Haloarchaeal diversity at 23, 121, and 419 MYA salts. *Geobiology* 7: 515-523.
- Parks, D.H. and **Beiko, R.G.** (2010) Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 26: 715-721.
- Parks, D.H., Porter, M., Churcher, S., Wang, S., **Blouin, C.**, Whalley, J., Brooks, S., **Beiko, R.G.** (2009) GenGIS: A geospatial information system for genomic data. *Genome Research* 19: 1896-1904.
- Perry, S.C. and **Beiko, R.G.** (2010) Distinguishing microbial genome fragments based on their composition: evolutionary and comparative genomic perspectives. *Genome Biology and Evolution* 2: 117-131.
- Rowe, D.L., Dunn, K.A., Adkins, R.M. and R.L. Honeycutt. (2010) Molecular clock keeps dispersal hypothesis afloat: evidence for trans-Atlantic rafting by rodents. *Journal of Biogeography* 37: 305-324.
- Saffo, M.B., McCoy, A.M., Rieken, C. and **Slamovits, C.H.** (2010) Nephromyces, a beneficial apicomplexan symbiont in marine animals. *Proc.Natl. Acad. Sci. USA* 107: 16190-16195.
- Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartman, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: open source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75: 7537-7541.
- Sebé-Pedrós, A., **Roger, A.J.**, Lang, F.B., King, N. and Ruiz-Trillo, I. (2010) From the cover: Ancient origin of the integrin-mediated adhesion and signaling machinery. *Proc. Natl. Acad. Sci USA* 107: 10142-10147.
- Silver, T.D., Moore, C.E., and **Archibald, J.M.** (2010) Nucleomorph ribosomal DNA and telomere dynamics in chlorarachniophyte algae. *Journal of Eukaryotic Microbiology* (in press; accepted August 24)
- Slamovits, C.H.** and Keeling, P.J. (2009) Evolution of ultrasmall spliceosomal introns in highly reduced nuclear genomes. *Molecular Biology and Evolution* 26: 1699-1705.
- Slater, F.R., Johnson, C.R., Blackall, L.L., **Beiko, R.G.** and Bond, P.L. (2010) Monitoring associations between clade-level variation, overall community structure and ecosystem function in enhanced biological phosphorus removal (EBPR) systems using terminal-restriction fragment length polymorphism (T-RFLP). *Water Research* [July 16 – Epub ahead of print, PMID: 20701946].
- Susko, E.** (2010) First-order correct bootstrap support adjustments for splits that allow hypothesis testing when using maximum likelihood estimation. *Molecular Biology and Evolution* 27: 1621-1629.

- Tanifuji, G. (2009) Have Diatoms ever possessed Prasinophyte-derived plastid? (Directly translated from Japanese). *Japan Journal of Phycology* 57: 134-136.
- Tanifuji, G. and **Archibald, J.M.** (2010) Actin gene family dynamics in cryptomonads and red algae. *Journal of Molecular Evolution* 71: 169-179.
- Tanifuji, G., Onodera, N.T. and Hara, Y. (2010) Nucleomorph genome diversity and its phylogenetic implications in cryptomonad algae. *Phycology Research* 58: 230-237.
- Tsaousis, A.D., Gaston, D., Stechmann, A., Walker, P.B., Lithgow, T. and **Roger A.J.** (2010) A Functional Tom70 in the Human Parasite *Blastocystis* sp.: Implications for the Evolution of the Mitochondrial Import Apparatus. *Molecular Biology and Evolution* [Sept. 24 - Epub ahead of print]
- Tsaousis, A.D., Leger, M.M., Stairs, C.W. and **Roger A.J.** (2010) The biochemical adaptations of mitochondrion-related organelles of parasitic and free-living microbial eukaryotes to low oxygen environments. In *Anoxia: Paleontological Strategies and Evidence for Eukaryote Survival*, A. Altenbach, J. Bernhard & J. Seckbach (Eds.), Springer-Verlag (*in press*).
- Volkman, G., Murphy, P.W., Rowland, E.E., Cronan, J.E. Jr., Liu, X.Q., **Blouin, C.** and Byers, D.M. (2010) Intein-mediated cyclization of bacterial acyl carrier protein stabilizes its folded conformation but does not abolish function. *Journal of Biological Chemistry* 285: 8605-8614.
- Whidden, C., **Beiko, R.G.** and Zeh, N. (2010) Fast FPT algorithms for computing rooted agreement forests: Theory and experiments. In *Proceedings of the 9th International Symposium on Experimental Algorithms, SEA 2010. Lecture Notes in Computer Science* 6049: 141-153, Springer-Verlag.
- Wilson, S., **Simpson, A.G.B** and Lynn, D.H. (2010) Maintaining journal figure quality. *Journal of Eukaryotic Microbiology* 57: 285-293.
- Wu, J. and **Susko, E.** (2010) Rate-variation need not defeat phylogenetic inference through pairwise sequence comparisons. *Journal of Theoretical Biology* 263: 587-589.
- Wu, J. and **Susko, E.** (2009) General heterotachy and distance method adjustments. *Molecular Biology and Evolution* 26: 2689-2697.
- Wang, H.C., **Susko, E.** and **Roger, A.J.** (2010) Fast statistical tests for detecting heterotachy in protein evolution. *Molecular Biology and Evolution* (*under review*).
- Zou, L., Field, C., **Susko, E.** and **Roger, A.J.** (2010). The Barry and Hartigan general Markov model suffers from statistical non-identifiability. *Systematic Biology* (*accepted for publication*).

VII. PRESENTATIONS BY CGEB MEMBERS (total of 30)

Archibald, J.M. *Endosymbiosis and genome mosaicism in cryptophyte and chlorarachniophyte algae*, 11th International Colloquium on Endocytobiology and Symbiosis, Tromsø, Norway, August 2010; Invited presentation

Archibald, J.M. *How do microbes rule the world?* ‘The Next Big Question’ National Tour, sponsored by the Canadian Institute for Advanced Research, Edmonton, May 11, 2010; Invited presentation

Archibald, J.M. *You are what you eat: endosymbiosis and genome evolution in unicellular eukaryotes*, London Natural History Museum, Seminar Series, London, UK, November 2009; Invited seminar

Archibald, J.M. *Bonsai genomics: origin and evolution of reduced endosymbiotic nuclear genomes*, University of Exeter, School of Biosciences Seminar Series, Exeter, UK, November 2009; Invited seminar

Archibald, J.M. *You are what you eat: endosymbiosis and genome evolution in cryptophyte and chlorarachniophyte algae*, Mount Allison University, Biochemistry Department Seminar Series, Sackville, October 2009; Invited seminar

Beiko, R.G. *SeqMonitor and GenGIS: Tools for collection and interactive visualization of molecular diversity*, University of Guelph, August 12, 2010; Invited presentation

Beiko, R.G. *It's a phylogenetic network and everyone's invited*, Beyond the Tree of Life Meeting, London, UK, July 2010; Invited presentation

Beiko, R.G. *Geographic and temporal analysis of genomes and metagenomes*, Canadian Institute for Advanced Research, Big DATA Meeting, Vancouver, BC, June 2010

Bielawski, J.P. *Codon substitution models and analysis of natural selection pressure in protein-coding genes*, Workshop on Molecular Evolution, Europe, January 2010; Invited presentation

Bielawski, J.P. *Searching for functional divergence in genes and genomes*, Killam Professor Award Seminar, Department of Biology, Dalhousie University, April 2010; Invited seminar

Bielawski, J.P. *Modelling protein evolution*, Workshop on Molecular Evolution, USA, August 2010; Invited presentation

Bielawski, J.P. *Teaching molecular evolution to advanced students*, Marine Biological Laboratories, Woods Hole, MA, September 2010; Invited round-table participant

Blouin, C. (with David N. Langelaan, M. Wiczorek and J.K. Rainey). *MC-HELAN: A Monte Carlo method for helix and kink characterization in proteins*, CSBMCB 2010, Banff, Alberta, April 15-18, 2010; Poster presentation

- Doolittle, W.F.** *Microbiology as the last refuge of typological thinking*, Current Issues in Darwinian Theory Workshop, Dalhousie University, Halifax, October 15, 2009; Invited presentation
- Doolittle, W.F.** *The Tree of Life: what does it mean?* Mount Allison University, Sackville, October 26, 2009; Invited seminar
- Doolittle, W.F.** *Irremediable complexity*, Biochemistry and Molecular Biology, Dalhousie University, November 26, 2009; Departmental seminar
- Doolittle, W.F.** *Irremediable complexity*, Department of Botany, University of British Columbia, Vancouver, BC, April 7, 2010; Invited seminar
- Doolittle, W.F.** *The Tree of Life, what it means and how it stands*, University of Alberta, Edmonton, AB, April 9, 2010; Invited Distinguished Lecture
- Doolittle, W.F.** *Was Darwin wrong (about the Tree of Life)?* International Behavioural and Neural Genetics Society, Dalhousie University, May 12, 2010; Invited Public Lecture
- Doolittle, W.F.** *Microbial systematics and the Tree of Life*, 2010 Annual Meeting of Canadian Society of Microbiologists, McMaster University, Hamilton, June 14, 2010; Invited keynote speaker
- Doolittle, W.F.** *Craig Venter's latest stunt: Thought Experiments in Synthetic Biology*, Situating Science Workshop on Science Without Data: The Role of Thought Experiments in Empirical Investigations, Halifax, June 18-19, 2010; Invited presentation
- Gray, M.W.** *Two Views of Mitochondrial Evolution: Genomics vs. Proteomics*, Department of Biology, University of New Brunswick, Fredericton, November 13, 2009; Invited seminar presentation
- Roger, A.J.** *Is a 'vertical' signal in deep prokaryote evolution recoverable given the 'horizontal' noise?* New Zealand Annual Phylogenetics Meeting (Doom 10), Ruapehu, New Zealand, February 10, 2010
- Roger, A.J.** *Early eukaryotic evolution: phylogenomics, rocks, grease and clocks*. Discussion meeting on "Dating early events in Earth's history," NASA Astrobiology Institute at UCLA, Los Angeles, CA, March 17-20, 2010; Invited presentation
- Roger, A.J.** *Phylogenomics - dataset construction and analyses*, CIFAR-IMB Program – Workshop on Eukaryotic Genome Annotation, Dalhousie University, June 2, 2010; Invited presentation
- Roger, A.J.** *Timing deep eukaryote phylogeny and evolution: Reconciling molecules, rocks, grease and clocks*, International Society for Evolutionary Protistology XVIII, Kanazawa, Japan, July 2010; Invited presentation

Simpson, A.G.B. *Protozoa from extraordinary environments*, Charles University, Prague, Czech Republic, May 2010; Invited seminar

Simpson A.G.B. and Park, J.S. *Flagellates from extraordinary environments*, Joint Meeting of the International Society of Protistologists and the British Society of Protist Biology, Canterbury, UK, July 2010; Invited presentation

Simpson A.G.B. *On the diversity of halophilic protozoa*, 18th meeting of the International Society of Evolutionary Protistology (ISEP), Kanazawa, Japan, July 2010; Invited presentation

Slamovits, C. *Eukaryotic genomes: diversity, function and architecture*, Institute of Biochemical Investigations “Dr. Luis F. LeLoir”, Buenos Aires, Argentina, December 15, 2009

VIII. OTHER RESEARCH CONTRIBUTIONS FROM CGEB TRAINEES

This section includes posters, talks, and seminar presentations (**total of 45**) by CGEB trainees at universities, national and international meetings, conferences and workshops. Travel and attendance was sponsored in part by CGEB (Tula) funds.

Note: CGEB trainees are underlined and the presenter's name is marked with an asterisk.

A. Seminars and Conference Presentations

Alfaro, J.* Gaston, D., Susko, E. and Roger, A. *Applications of structurally constrained models of protein evolution to problems in structural bioinformatics*, 3DSIG 2010 (ISMB Satellite Meeting), Boston, MA, July 9-10, 2010 (poster)

***Received Graduate Student Conference Travel Grant from FGS**

Eveleigh, R.J.M.* Archibald, J.M. and Beiko, R. *Being Aquifex aeolicus: Untangling a hyperthermophiles checkered past*, The Robert Cedergren Bioinformatics Colloquium, Montreal, Quebec, November 5-6, 2009 (poster)

Gaston, D.* and Roger, A.J. *Evaluating the performance of Functional Divergence prediction methods in protein evolution*, 2010 Informatics for Phylogenetics, Evolution and Biodiversity (iEvoBio 2010) International Conference, Portland, OR, June 2010 (oral presentation). ***Received Graduate Student Conference Travel Grant from FGS**

Gaston, D.* and Roger, A.J. *CBOrg: Comparative BLAST for Organelles*, Evolution 2010 International Conference, Portland, OR, June 2010 (poster)

Gawryluk, R.M.* *A proteomic investigation of the Acanthamoeba castellanii mitochondrial electron transport chain*, Department of Biochemistry & Molecular Biology, Dalhousie University, October 2010 (seminar)

Heiss, A.A.* Walker, G. and Simpson, A.G.B. *The cytoskeleton of Apusozoa, and its implications for deep eukaryotic cell evolution*, International Society for Evolutionary Protistology (ISEP) Meeting XVIII, July 2-7, 2010, Kanazawa, Japan (oral presentation)

***Received ISEP Student Travel Award**

Heiss, A.A.* Walker, G. and Simpson, A.G.B. *The morphology of the Apusozoa, a possible key to understanding the evolution of eukaryotes*, 3rd Annual Dr. Patrick Lett Symposium -- graduate student conference, Biology Department, Dalhousie University, May 6, 2010 (oral presentation)

Hopkins, J.F.* Spencer, D.F., Gray, M.W. and Archibald, J.M. *Proteomics reveals complex evolution of plastid- and nucleomorph-targeted proteins in the chlorarachniophyte Bigelowiella*, SMBE Meeting, Lyon, France, July 4-8, 2010 (oral presentation)

Kim, E.* *Evolution of eukaryotic microbes: exploring novel diversity and biotic associations*, Department of Biochemistry and Molecular Biology, Dalhousie University, September 30, 2010 (seminar)

Kim, E.* Park J.S., Simpson, A.G.B., Matsunaga, S., Watanabe, M., Murakami, A., Sommerfeld, K., Onodera, N.T. and Archibald, J.M. *Complex array of endobionts in Petalomonas sphagnophila, a large heterotrophic euglenid protist from Sphagnum-dominated peatlands*, International Society of Microbial Ecology, Seattle, WA, August 21, 2010 (poster)

Kim, E.* and Archibald, J.M. *Evolution and ecology of uncultured eukaryotic microbes*, CIFAR – Integrated Microbial Biodiversity Program Annual Meeting, Seattle, WA, August 24, 2010 (oral presentation)

Kolisko, M.* *Phylogenomic analysis of short-branching relatives of diplomonads*. CIFAR – Integrated Microbial Biodiversity Program Annual Meeting, Seattle, WA, August 2010 (oral presentation)

Kolisko, M.* and Roger, A.J. *Phylogenomics—dataset construction and analyses*. CIFAR, Workshop on Eukaryotic Genome Annotation, Dalhousie University, Halifax, June 1-3, 2010 (oral presentation)

Leger, M.* Gawryluk, R.M.R., Gray, M.W. and Roger, A.J. *Anaerobic metabolism enzymes in Acanthamoeba castellanii*, 18th Meeting of the International Society for Evolutionary Protistology (ISEP), Kanazawa, Japan, July 2-7 2010 (oral presentation).
***Received Graduate Student Conference Travel Grant from FGS**

Leger, M.* and Roger, A.J. *The jakobid Andaluca incarcerate has a hydrogenosome; implications for convergence on anaerobic metabolic pathways in eukaryotes*, Canadian Institute for Advanced Research-Integrated Microbial Biodiversity Program meeting, Seattle, WA, August 19-22, 2010 (oral presentation).

MacDonald, N.J.* and Beiko, R.G. *Discovering genotype-phenotype maps with association rule mining*, SMBE 2010, Lyon, France (poster).

MacDonald, N.J.* and Beiko, R.G. *Mitigating the effect of shared ancestry in genotype-phenotype association problems with conditional mutual information*, Robert Cedergren Colloquium, Montreal, Quebec, November 2009 (oral presentation & poster).

MacDonald, N.J.* and Beiko, R.G. *Increasing the generalizability of genotype-phenotype classifiers using conditional mutual information*, Atlantic Provinces Council on the Sciences (APICS), Halifax, NS, October 2009 (oral presentation)

Malik S.-B.* and Carlton, J.M. *Phylogeny, paralogy and sex in trichomonads*, New York Area Microbial Evolution Group (monthly meeting), American Museum of Natural History, New York, November 17, 2009 (seminar)

Malik, S.-B.* Bilic, I., Brochu, C., Conrad, M.D., Jelcic, M.J., Sotolongo, K., Sullivan, S.A., Venturini, E., Winn, A.S., Zavadil, J., Logsdon, J.M., Hess, M. and Carlton, J.M. *Comparison of Parabasalid protein-coding genes for phylogeny and sex*, Canadian Institute for Advanced Research-Integrated Microbial Biodiversity Program Annual Meeting, Seattle, WA, August 21, 2010 (poster)

Malik S.-B.* Winn, A.S., Venturini, E., Sotolongo, K., Jelcic, M.J., Conrad, M.D., Sullivan, S.A. and Carlton, J.M. *Phylogeny, paralogy and sex in trichomonads*, Parasite Discussion Group, Department of Medical Parasitology, New York University Langone Medical Center, NY, January 27, 2010 (seminar)

Maruyama, S.* *Genome evolution in lower plants and protists, and the origin of plastids*, Botanical Society of Japan (BSJ), Sept. 10, 2010 (oral presentation)

Moore, C.E.* and Archibald, J.M. *Small genomes, big questions: genome reduction as revealed by the large nucleomorph genome of the cryptophyte alga Chroomonas*, Society for Molecular Biology and Evolution (SMBE) annual meeting, Lyon, France, July 4-8, 2010 (oral presentation). ***Received SMBE Graduate Student Travel Award.**

Onodera, N., Tanifuji, G.*, Donaher, N., Malfatti, S., Chain, P., Hara, Y. and Archibald, J. *The complete plastid genome sequence of the secondarily non-photosynthetic alga Cryptomonas paramecium: Reduction, compaction, and accelerated evolutionary rate*, International Society for Evolutionary Protistology, Kanazawa, Japan, July 2-7, 2010 (poster)

Parks, D.H.* and Beiko, R.G. *Investigating viral and microbial biogeography using GenGIS*, International Society for Microbial Ecology (ISME), Seattle, WA, August 2010 (poster)

Parks, D.H.* and Beiko, R.G. *GenGIS: a visualization framework for exploring georeferenced genomic and metagenomic data sets*, EMBO Workshop on Visualizing Biological Data (VizBi), Heidelberg, Germany, March 2010 (poster)

Parks, D.H.* MacDonald, N.J. and Beiko, R.G. *A rapid data acquisition pipeline for visualizing the 2009 A/H1N1 influenza pandemic*, Atlantic Provinces Council on the Sciences (APICS), Halifax, NS, October 2009 (oral presentation)

Porter, M.* *GenGIS: 3D interactive visualizations of biodiversity*, iEvoBio 2010 Conference, Portland, OR, June 29, 2010 (oral presentation)
***Awarded 2nd place prize (for GenGIS software) in iEvoBio's Visualization Challenge** – award co-shared with D.H. Parks

Porter, M.* and Hajibabaei, M. *iBarcode-nextgen: Tools for next-generation sequence analysis*, iEvoBio 2010 Conference, Portland, OR, June 29, 2010 (oral presentation)

Simpson, A.G.B. and Park, J.S.* *On the diversity of halophilic protozoa*, International Society for Evolutionary Protistology, XVIII, Kanazawa, Japan, July 2010 (oral presentation)

Stairs, C.W.*, Roger, A.J. and Hampl, V. *The origin of enzymes of anaerobic pyruvate metabolism in protists: the case of pyruvate formate lyase and its activating enzyme*, International Society for Evolutionary Protistology, XVIII, Kanazawa, Japan, July 2010 (oral presentation) ***Received ISEP Student Travel Award**

Stevens, G.C., Tsaousis, A.D.* and Roger, A.J. *Strategies of pyruvate decarboxylation of the mitochondrion-related organelle of Blastocystis sp.*, ISEP XVIII meeting, Kanazawa, Japan, July 2-7, 2010 (poster)

Takishita, K.*, M. Kolisko, H. Komatsuzaki, A. Yabuki, I. Cepicka, P. Smejkalova, J.D. Silberman, Y. Inagaki, T. Hashimoto, A.J. Roger and A.G.B.Simpson. *Multigene phylogenies of diverse Carpediemonas-like organisms identify the closest relatives of 'amitochondriate' diplomonads and retortamonads*, ISEP XVIII meeting, Kanazawa, Japan, July 2-7, 2010 (oral presentation)

Tanifuji, G.* Onodera, N., Travis, W., Donaher, N., Dlutek, M. and Archibald, J. *Complete nucleomorph genome sequence of the non-photosynthetic alga Cryptomonas paramecium reveals unexpected similarities between the cryptophyte and chlorarachniophyte nucleomorph genomes*, International Society for Evolutionary Protistology, Kanazawa, Japan, July 2-7, 2010 (oral presentation)

***Received ISOP/ISEP Travel Award**

Tsaousis, A.D.* *The Mysterious Organelle of Blastocystis sp.: Implications for the Evolution of Mitochondria*, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS, May 27, 2010 (seminar)

Tsaousis, A.D.* and Roger, A.J. *The unique mitochondrion related organelle of Blastocystis*, 21st Annual Molecular Parasitology Meeting, MBL, Woods Hole, MA, September 12-16, 2010 (oral presentation)

Wang, H.C.* *An improved w statistics for covarion and heterotachy tests*, 2nd Annual Atlantic Mathematical Biology Workshop, University of New Brunswick, Fredericton, NB, June 7-8, 2010 (oral presentation)

Whidden, C.* and Zeh, N. *A unifying view on approximation and FPT of agreement forests*, APICS 2009 Computer Science Conference, Dalhousie University, Halifax, NS, October 23-24, 2009 (oral presentation)

***Received Best Graduate Student Presentation Award**

Whidden, C., Beiko, R.G. and Zeh, N.* *Fast FPT algorithms for computing rooted agreement forests: Theory and experiments*, SEA 2010, Ischia Island (Napoli), Italy, May 20-22, 2010 (oral presentation).

Wong, D.* and Beiko, R.G. *A Phylogenomic and Metabolic Analysis of EBPR Community Metagenomes*, Robert Cedergren Bioinformatics Colloquium, Montreal, QC, November 2009 (oral presentation)

Wong, D.* and Beiko, R.G. *Phylogenomic and Metabolic Contrasts Between Two Enhanced Biological Phosphorous Removal Communities*, 13th International Symposium on Microbial Ecology (ISME), Seattle, WA, August 2-27, 2010 (poster presentation)

B. Invited Presentations from CGEB Trainees

Malik S.-B.* Pightling, A.W., Logsdon, J.M. Jr., Venturini, E., Carlton, J.M., Leander, B.S. and Slamovits, C.H. *Dear Protists: Got Meiosis? How do you maintain your (genome) integrity?* Canadian Institute for Advanced Research, Junior Fellow Academy Annual Meeting, Toronto, June 7, 2010 (invited presentation)

Parks, D.H.* and Beiko, R.G. *Using β -diversity to study the influence of geographic proximity and habitat on the composition of marine microbial communities*, 95th Ecological Society of America Annual Meeting, Pittsburgh, PA, August 2010 (invited presentation)

Tsaousis, A.D.* *Insights on the function of the mitochondrion-related organelle of Blastocystis sp.*, Department of Biomedical Chemistry, University of Tokyo, Tokyo, Japan, July 16, 2010 (invited lecture)

Tsaousis, A.D.* *Insights on the function of the mitochondrion-related organelle of Blastocystis sp.*, ISEP XVIII Meeting, Kanazawa, Japan, July 2-7, 2010 (invited presentation)

IX. COLLABORATIVE INTERACTIONS BETWEEN CGEB MEMBERS

In the last year, collaborative interactions amongst CGEB labs have flourished. Evidence of this can be seen in the 7 publications that were co-authored by two or more CGEB members and their trainees (pp. 10-15). Below a short summary is provided of the collaborations that are ongoing within CGEB:

Archibald and Gray are working on the analysis and annotation of cryptophyte nuclear and organellar genome data coming from the Joint Genome Institute. They are co-PIs on an NSERC Strategic Opportunities Grant to sponsor this work and co-supervise Tula-sponsored postdoc Dr. Goro Tanifuji.

Beiko and Archibald are developing phylogenomic methods for analysing eukaryote and prokaryote genomic data to be integrated into the MOA database. They co-supervise M.Sc. student Rob Eveleigh.

Bielawski and CGEB associate Hong Gu continue to collaborate on a new statistical framework for modeling complex microbial phenotypes as networks. This research is expected to be greatly enhanced by their co-supervision of a postdoc who will be funded by the CIHR Emerging Team Grant recently awarded to Doolittle, Bielawski and Beiko and who will have wide interaction and collaboration with CGEB members as part of the emerging team.

Bielawski and Susko are collaborating on the development of codon models with theoretically justifiable interpretations with respect to population genetic parameters such as selection coefficients and effective population size. As part of this work they are co-supervising the M.Sc. program of Joey Mingrone.

Blouin and Beiko have been involved in the development of methods (e.g. the software tool GenGIS) that integrate geospatial data with genomic (and metagenomic) data.

Doolittle, Bielawski and Beiko teamed up to write a CIHR Emerging Team Grant entitled “Modeling and mapping microbial diversity and function with marker genes, genomes and metagenomes” that was submitted in the April 2010 CIHR Canadian Microbiome Initiative competition. This application was ranked first in its competition, and fully funded (Sept. 2010-Aug. 2013). This underwrites the costs of three years of collaborative research.

Roger and Beiko began a collaboration last year to determine whether an ancient vertical phylogenetic signal exists for core proteins in the domain Bacteria or if it has been completely erased as some have claimed. This project is still simmering on the backburner and will benefit from the new algorithms being developed by CGEB trainee Chris Whidden.

Roger and Gray groups are working together to elucidate the origin and function of enzymes of anaerobic energy metabolism in the aerobic amoeba *Acanthamoeba castellanii*.

Roger and Simpson are collaborating on understanding the phylogeny and evolution of the most enigmatic eukaryote supergroup – the Excavata by EST surveys and phylogenomic analysis. They co-supervise Ph.D. student Martin Kolisko.

Roger and Simpson have begun a collaboration to examine molecular evolution in very halophilic (salt-loving) protozoa, and to use EST-based surveys to shed light on adaptation extreme salinity. They co-supervise Ph.D. student Tommy Harding, and the work also involves Tula-sponsored postdoc Dr. Jong Soo Park.

Susko and Doolittle have collaborated over the last few years in developing methods with Eric Bapteste (former PDF) to probe the extent of lateral gene transfer in prokaryotic genome evolution.

Susko, Roger, and Blouin continue to work on the development of methods for curating protein alignments and site-independent models of protein evolution that take into account structural information.

Susko and Roger have been developing tools for phylogenomic analysis, ancestral sequence estimation, sophisticated models of protein evolution methods for assessing the quality of EST data and theoretical contributions to the understanding of molecular evolution. They work together with Tula-sponsored postdoc Dr. Huaichun Wang (supervised by Susko and Roger), PhD students Dan Gaston (supervised by Roger) and Liwen Zou (co-supervised by Susko, Field and Roger), and M.Sc. student Javier Alfaro (supervised by Roger).

X. CGEB DEVELOPED SOFTWARE

2009-2010:

Program name: *FunDi*

Authors: Dan Gaston, Ed Susko and Andrew Roger

Description: A maximum-likelihood based phylogenetic mixture model for prediction of functionally divergent protein residues. Implemented in Perl along with Perl libraries for interfacing with RAxML, QmmRAxML, and PUZZLE.

Website: <http://cortana.biochem.dal.ca/software>

Program: *GenGIS - A Geospatial Information System for Genetic Data*

Authors: Parks, D.H., Porter, M., Churcher, S., Wang, S., Blouin, C., Whalley, J., Brooks, S. and Beiko, R.G.

Description: *GenGIS* is a bioinformatics application that allows users to combine digital map data with information about biological sequences collected from the environment. *GenGIS* provides a 3D graphical interface in which the user can navigate and explore the data, as well as a Python interface that allows easy scripting of statistical analyses using the Rpy libraries.

Website: http://kiwi.cs.dal.ca/GenGIS/Main_Page

Program name: *PICA*

Authors: Norman MacDonald and Robert Beiko

Description: Phenotype Investigation with Classification Algorithms (PICA) is a Python framework for testing genotype-phenotype association algorithms.

Website: <http://kiwi.research.cs.dal.ca/Software/PICA>

Program name: *rSPR*

Authors: Chris Whidden, Norbert Zeh and Robert Beiko

Description: rSPR is a software package for calculating rooted subtree-prune-and-regraft distances and rooted agreement forests. Please note that this page is under construction.

Website: <http://kiwi.cs.dal.ca/Software/RSPR>

Program name: *STAMP*

Authors: Donovan Parks and Robert Beiko

Description: STAMP (Statistical Analysis of Metagenomic Profiles) is a software package for analyzing metagenomic profiles (e.g., a phylogenetic profile indicating the number of marker genes assigned to different taxonomic units or a functional profile indicating the number of sequences assigned to different biological subsystems or pathways) that promotes 'best practices' in choosing appropriate statistical techniques and reporting results. It encourages the use of effect sizes and confidence intervals in assessing biological importance. A user friendly, graphical interface permits easy exploration of statistical results and generation of publication quality plots for inferring the biological relevance of features in a metagenomic profile. STAMP is open source, extensible via a plugin framework, and available for all major platforms.

Website: <http://kiwi.research.cs.dal.ca/Software/STAMP>

Program name: *aBP*

Authors: Ed Susko

Description: A software tool implementing first-order correct bootstrap support adjustments for splits that allow hypothesis testing when using maximum likelihood estimation.

Website: <http://www.mathstat.dal.ca/~tsusko/>

2007-2009:

Program name: *minmax-chisq*

Authors: Ed Susko and Andrew Roger

Description: C language source for some of the methods described in Susko, E. and Roger, A.J. (2007). On reduced amino acid alphabets for phylogenetic inference. *Mol. Biol. Evol.* 24: 2139-2150.

Website: <http://www.mathstat.dal.ca/~tsusko/>

Program name: *PROCOV-SPR (Protein COVarion analysis)*

Authors: Huaichun Wang, Ed Susko and Andrew Roger

Description: A maximum likelihood phylogenetic inference program that implements a variety of rates-across-sites and covarion models and allows tree-searching using the SPR function.

Website: <http://www.mathstat.dal.ca/~hewang/Procov/>

Program name: *QmmRAxML (Q-matrix mixture RAxML)*

Authors: Huaichun Wang, Ed Susko and Andrew Roger

Description: A maximum likelihood phylogenetic inference program that implements a mixture model with any number of rate (Q) matrices and frequency vectors. Captures the 'site specific' nature of protein evolution.

Website: <http://www.mathstat.dal.ca/~hewang/QmmRAxML/>

Program name: *CONCATERPILLAR*

Authors: Jessica Leigh, Ed Susko and Andrew Roger

Description: A program that implements hierarchical likelihood ratio test and a clustering algorithm for testing phylogenetic congruence between partitions of a phylogenomic dataset.

Website: <http://rogerlab.biochemistryandmolecularbiology.dal.ca/Software/Software.htm>

Program name: *Barrel-o-Monkeys*

Authors: Jessica Leigh and Andrew Roger

Description: A series of Python scripts that automate repetitive tasks in bioinformatics that are so simple, but annoying, that a monkey could do them.

Website: <http://rogerlab.biochemistryandmolecularbiology.dal.ca/Software/Software.htm>

Program name: *RAADdet*

Authors: Jessica Leigh, Ed Susko and Andrew Roger

Description: A program that implements novel methods for adjusting for rates-across-sites in an amino acid Log-Det distance calculation.

Website: <http://rogerlab.biochemistryandmolecularbiology.dal.ca/Software/Software.htm>

Program name: *CBOrg*

Authors: Daniel Gaston and Andrew Roger

Description: A tool that does comparative BLAST analyses on a set of input sequences to determine if their gene products are targeted to mitochondria or hydrogenosomes versus elsewhere in the cell.

Website: <http://rogerlab.biochemistryandmolecularbiology.dal.ca/Software/Software.htm>

Program: *Radié*

Authors: J. Whalley, S. Brooks and R.G. Beiko

Description: Radié is a Java tool that allows characters to be visualized against the background of a phylogenetic tree.

Website: <http://kiwi.cs.dal.ca/~beiko/software-and-data/radie>

Program: *Codeml FE*

Authors: Bielawski research group

Description: A modification of the program codeml (Yang, 1997) that implements 11 new fixed effect models of codon evolution.

Website: www.bielawski.info

Program: *LiBaC (Likelihood Based Clustering)*

Authors: Bielawski research group

Description: Implements a novel statistical technique to group sites according to similarities in the underlying process of molecular evolution. *LiBaC* is a generalization of the statistical technique called “model based clustering” to Markov models of codon evolution.

Website: www.bielawski.info

Program: *MANUEL*

Authors: Christian Blouin, Scott Perry, A. Lavell, Ed Susko and Andrew Roger

Description: An SVM-based multiple sequence annotation offered through a web interface.

Website: <http://fester.cs.dal.ca/manuel/>

Programs for cross validation on codon models of evolution

Authors: Joey Mingrone and Joseph Bielawski

Description: Implemented a suite of programs for using cross validation as a method to assess model based classification of individual codon sites to different groups based on the strength of natural selection pressure.

Website: <http://awarnach.mathstat.dal.ca/repos>

Programs for bootstrapping codon models of evolution

Authors: Joey Mingrone and Joseph Bielawski

Description: Employs bootstrapping as a method to assess model based classification of individual codon sites to different groups based on the strength of natural selection pressure. This work is in progress; software is made available at <http://awarnach.mathstat.dal.ca/repos> as it is being developed.

Programs and scripts for automated analysis of genomic data

Authors: Bielawski research group

Description: A suite of over 50 Perl programs for bioinformatic analysis. The newest programs allow automation of many methods for genome scale computational analysis. We have established a web site for public access to some of these programs (24 of the most commonly used programs) and supporting documentation.

Program: *spr distance*

Authors: Glenn Hickey, Frank Dehne, Andrew Rau-Chaplin and Christian Blouin

Description: An FTP-like algorithm to evaluate the length of the shortest SPR path between two phylogenetic trees.

Website: <http://peta.cs.dal.ca/blouinlab/node/7>

XI. CURRENT EXTERNAL RESEARCH GRANTS AND FUNDING

Collectively, CGEB faculty members received [~\$1,973,208] from external research grants and funding in 2009-2010. Grants involving more than one CGEB faculty (as co-applicants) were only included (i.e. counted) for the primary P.I. for the above total.

J.M. Archibald:

- 2009-2013 **NSERC Discovery Grant** —*Genome and proteome evolution in nucleomorph-containing algae*: \$34,000/year
- 2008-2013 **Canadian Institutes of Health Research – New Investigator Award**: 5-year salary award: \$60,000/year
- 2008-2010 **NSERC Special Research Opportunities Program Grant**—*Impact of secondary endosymbiosis on eukaryotic genome evolution*: \$113,050 (Year 1); \$106,050 (Year 2)
- 2008-2010 **CIHR & NS/CIHR Regional Partnership Program Operating Grant**—*The causes and consequences of genome reduction in eukaryotes*: \$100,000 (Year 1); \$76,615 (Year 2)
- 2007-2012 **Tula Foundation**, CGEB Molecular Biology Research Fellowship (funding for postdoctoral fellowship plus research expenses); total \$64,000/year for 5 years
- 2007-2012 **Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity**—research support funds: \$35,000/year

R. Beiko:

- 2010-2013 **Canadian Institutes of Health Research - Emerging Team Grant: Canadian Microbiome Initiative**: *Modeling and mapping microbial diversity and function with marker genes, genomes and metagenomes*: \$223,000/year (with co-PIs: W.F. Doolittle, J. Bielawski & M. Ereshefsky – only counted for Doolittle in above total)
- 2009-2012 **Genome Canada, Competition in Applied Genomics Research – BEEM**: *Bioproducts and enzymes from environmental metagenomes*: ~ \$10,000,000 over 3 years (8 co-applicants with E. Edwards, P.I. – not included in above total)
- 2008-2013 **Tula Foundation** – funding for PhD student traineeship plus research allowance; total \$27,500/year for 5 years
- 2007-2012 **NSERC - Discovery Grant**, *New computational methods for metagenomics*: total \$115,000 over 4 years
- 2007-2012 **Canada Research Chairs (CIHR) – Canada Research Chair (Tier II) in Bioinformatics**: total \$500,000
- 2007-2009 **Australian Research Council – Discovery Grant** – *Metagenomics and the genetic basis of ecology and evolution of communities*: \$400,000 (co-applicant with L.L. Blackall, P.I. & 2 others – not included in above total)
- 2006-2009 **Genome Atlantic – Startup Grant**: *Computational analysis of genomic and metagenomic data*: total \$224,000 over 3 years

J.P. Bielawski:

- 2010-2013 **Canadian Institutes of Health Research - Emerging Team Grant: Canadian Microbiome Initiative:** *Modeling and mapping microbial diversity and function with marker genes, genomes and metagenomes:* \$223,000/year
(with co-PIs: W.F. Doolittle, R. Beiko & M. Ereshefsky – only counted for Doolittle in above total)
- 2009-2013 **NSERC – Discovery Grant:** *Evolutionary analysis of protein diversification: model improvement, assessment and application to real data:* total \$145,000 over 4 years
- 2007-2012 **Tula Foundation** – funding for CGEB postdoctoral fellowship plus research allowance; total \$49,000/year for 5 years

C. Blouin:

- 2010-2015 **NSERC – Individual Discovery Grant:** *Exploring the landscape of phylogenies:* \$24,000/year
- 2005-2010 **NSERC – Individual Discovery Grant:** *Phylogenetics and protein evolution:* \$23,700/year
- 2009-2014 **CFI (Infrastructure Operating Funds):** *A high performance computational platform for bioinformatics:* \$12,000/year
- 2008-2013 **Tula Foundation** – funding for PhD student traineeship plus research allowance; total \$27,500/year for 5 years

W.F. Doolittle:

- 2010-2013 **Canadian Institutes of Health Research - Emerging Team Grant: Canadian Microbiome Initiative:** *Modeling and mapping microbial diversity and function with marker genes, genomes and metagenomes;* \$223,000/year
(with co-PIs: J.P. Bielawski, R.G. Beiko and M. Ereshefsky)
- 2008-2010 **Tula Foundation,** CGEB Molecular Biology Research Fellowship (funding for postdoctoral fellowship plus research expenses); total \$64,000/year for 2.5 years

M.W. Gray:

- 2009-2011 **Tula Foundation,** CGEB Molecular Biology Research Fellowship (funding for postdoctoral fellowship plus research expenses); total \$64,000/year for 2.5 years

A.J. Roger:

- 2010-2017 **Canada Research Chairs Program (CIHR) – Canada Research Chair (Tier I) in Comparative Genomics and Evolutionary Bioinformatics:** \$200,000/year
(associated CFI application for \$301,664 for equipment purchase and renovation currently approved by CFI, but awaiting matching funding from NSRIT)
- 2008-2013 **Canadian Institutes of Health Research, Operating Grant – Major transitions in eukaryotic cell evolution:** \$719,381 over 5 years
- 2008-2013 **Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity—** research support funds: \$25,000/year

- 2007-2015 **Tula Foundation**, Centre for Comparative Genomics & Evolutionary Bioinformatics (CGEB): Administrator salary, Seminar Series and Student Travel funds: \$50,000/year for 8 years
- 2007-2012 **Tula Foundation**, CGEB Molecular Biology Research Fellowship (funding for postdoctoral fellowship plus research expenses); total \$64,000/year for 5 years
- 2005-2010 **Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant** – *Phylogenomic approaches to ancient relationships amongst eukaryotes*: \$41,300/year for 5 years

A.G.B. Simpson:

- 2009-2014 **NSERC – Individual Discovery Grant:** *Microbial eukaryote diversity and evolution in extraordinary environments*: \$34,000/year
- 2008-2013 **Tula Foundation**, CGEB Molecular Biology Research Fellowship (funding for postdoctoral fellowship plus research expenses); total \$64,000/year for 5 years
- 2008-2013 **CIHR Operating Grant** (*co-applicant with A.J. Roger, P.I.*: \$143,876/year – *only counted for A.J. Roger in above total*)
- 2007-2011 **Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity** - research support funds: \$23,000/year

C.H. Slamovits:

- 2010-2015 **NSERC Discovery Grant:** *Genomes of alveolate protists: structure, function and evolution*: \$27,000/year
- 2009-2014 **Canada Foundation for Innovation – Leaders Opportunity Fund:** *A laboratory for microbial parasite genetics and evolution*: total \$312,500
- 2009-2014 **Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity**— research support funds: \$25,000/year
- 2009-2010 **Dalhousie DMRF Capital Equipment Grant:** total \$13,500

E. Susko:

- 2008-2013 **NSERC – Discovery Grant:** *Statistical evolutionary bioinformatics*: \$21,000/year
- 2008-2013 **Tula Foundation** – CGEB Postdoctoral Research Fellowship (funding for postdoctoral fellowship plus research expenses); total \$49,000/year for 5 years

XII. HONOURS/AWARDS/DISTINCTIONS

[includes those received from 2009-2010, ongoing salaried fellowships and scholarships, and lifetime fellowships and honorary memberships]

John Archibald:

- Fellow, Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program (2009-2012)
- New Investigator Award, Canadian Institutes of Health Research (5-year salary award, 2008-2013)

Robert Beiko:

- Canada Research Chair (Tier II) in Bioinformatics (2007-2012): \$100,000/year
- Recipient of Dalhousie Student Union Teaching Award (2009-2010)
- Award for GenGIS – Finalist in Biological Visualization, VizBi, Heidelberg (Feb. 2010)
- Awarded 2nd place prize (for *GenGIS* software) in iEvoBio's Visualization Challenge, Portland, Oregon (June, 2010)

Joseph Bielawski:

- Recipient of Dalhousie Faculty of Science Killam Prize (2009)

W. Ford Doolittle:

- Elected Member, The Norwegian Academy (2009-present)
- Institute Fellow, Canadian Institute for Advanced Research (2008-present)
- Elected Member, U.S. National Academy of Sciences (2002-present)
- Fellow, American Academy of Microbiology (1999-present)
- Fellow, The Royal Society of Canada (1991-present)
- Fellow, American Association for the Advancement of Science (1985-present)

Michael Gray:

- Elected Fellow, Royal Society of Canada, Academy of Sciences (1996-present)

Andrew Roger:

- Canada Research Chair (Tier I) in Comparative Genomics and Evolutionary Bioinformatics (2010-2017): \$200,000/year
- Recipient of the 2010 Seymour Hutner Award from the International Society of Protistologists
- President-Elect (2010-2012), International Society for Evolutionary Protistology (ISEP)
- Director, Centre for Comparative Genomics & Evolutionary Bioinformatics (CGEB), Dalhousie University (2008-2011)
- Fellow, Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program (2007-2012)

Alastair Simpson:

- Recipient of the Dalhousie Faculty of Science Killam Prize (2010)
- Fellow, Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program (2009-present)

XIII. SERVICE TO THE SCIENTIFIC COMMUNITY (OUTSIDE DALHOUSIE)

A. Journal Editorial Boards (2009-2010)

Collectively, the ten CGEB faculty members serve on the following Editorial Boards:

Archaea; Biology Direct; BMC Biology; BMC Evolutionary Biology; Canadian Journal of Statistics; Environmental Microbiology; Gene; Genome Biology; IUBMB Life; Journal of Eukaryotic Microbiology; Journal of Experimental Zoology, Series B; Journal of Molecular Evolution; Journal of Phycology; Molecular Biology and Evolution; Phycological Research; Proceedings of the US National Academy of Sciences; Protistology; Science; Systematic Biology; Trends in Microbiology

B. Manuscript Peer Review

As well, CGEB members routinely review manuscripts for the above and for :

Acta Protozoologica; Annals of Statistics; Annals of Applied Statistics; Archiv für Protistenkunde; Biochemica et Biophysica Acta; Bioessays; Bioinformatics; Bioinformatics and Biology Insights; Biology Letters; Biology and Philosophy; Biological Bulletin; Biometrics; Biotechniques; Botanica Marina; BMC Bioinformatics; BMC Genomics; BMC Structural Biology; Computational Statistics and Data Analysis; Copeia; Current Biology; Current Genetics; Current Opinion in Microbiology; Encyclopedia of Life Sciences; Eukaryotic Cell; Evolution; Evolutionary Biology Online; Evolutionary Bioinformatics; FEBS Letters; Genetics; Genetics Research; Genome; Genomics; Genome Biology and Evolution; Heredity; International Journal of Biochemistry and Cell Biology; International Journal for Parasitology; International Journal of Systematic and Evolutionary Microbiology; Journal of Biology; Journal of Biomedical Informatics; Journal of Cell Science; Journal of Computational Biology; Journal of Fish Biology; Journal of Multivariate Analysis; Journal of Plant Physiology; Journal of Structural Biology; Journal of Theoretical Biology; Microbiology; Molecular Ecology; Molecular Genetics and Genomics; Molecular Microbiology; Molecular Phylogenetics and Evolution; Nature; Nature Reviews Genetics; Nature Reviews Microbiology; New Phytologist; Nucleic Acids Research; Philosophical Transactions of the Royal Society, Series B: Biological Sciences; Physics Letters A; PLoS One; PLoS Biology; PLoS Genetics; Proceedings of the Royal Society of London; Proceedings of Mathematics of Evolution and Phylogenetics; Protein Science; Protist; Public Library of Science (Biology); Recomb-2005; RNA; Statistical Science; Statistica Sinica; Symbiosis; Theoretical Population Biology; Trends in Ecology and Evolution; Trends in Genetics; Zoology

C. National & International Service on Scientific Advisory Boards/Committees/Grant Panels:

- Associate Director, **Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity** (J.M. Archibald)
- Treasurer, **International Society for Molecular Biology and Evolution** (J.M. Archibald)
- Peer Review Committee Member, Doctoral Research Award A (DRA) and Masters Awards (CGA) **Canadian Institutes of Health Research** (J.M. Archibald)
- Peer Review Committee Member, Human Microbiome Catalyst Grants, **Canadian Institutes of Health Research** (J.M. Archibald)
- Member, Scientific Advisory Group, **Ontario Genomics Institute** (J.M. Archibald; R.G. Beiko and W.F. Doolittle)
- Grant Reviewer for: **Canada Research Chairs Program; Swiss National Foundation; German-Israeli Foundation; UK Medical Research Council** (R.G. Beiko)
- Grant Reviewer for: **NSERC; National Science Foundation (USA); Israeli Science Foundation** (R.G. Beiko and J.P. Bielawski)
- Member, Steering Committee for the Workshop on Molecular Evolution, **Marine Biological Laboratories (USA)** (J.P. Bielawski)
- Committee Member, **Atlantic Association for Research in Mathematical Science** (C. Blouin)
- Committee Member, **Canadian National Committee for CODATA** (C. Blouin)
- Chair, Scientific Advisory Board, **Genome Alberta Project: Metagenomics for greener production and extraction of hydrocarbon energy** (W.F. Doolittle)
- Member, Expert Panel on the State and Trends of Biodiversity Science in Canada, **Council of Canadian Academies** (W.F. Doolittle)
- Advisory Committee, **Canadian Institute for Advanced Research – Integrated Microbial Biodiversity Program** (W.F. Doolittle and M.W. Gray)
- Member, Science and Engineering Advisory Council, **Alberta Ingenuity Fund** (M.W. Gray)
- Grant Reviewer for: **National Science Foundation (USA); Canadian Institutes of Health Research** (A.J. Roger)
- Grant Reviewer for: **Czech Academy of Sciences** (A.G.B. Simpson)
- Grant Reviewer for: **NSERC; Marsden Fund Grants, New Zealand** (E. Susko)
- Consultant for **Department of Fisheries and Oceans** (E. Susko)

XIV. OTHER ACTIVITIES OF THE CENTRE

A. CGEB Seminar Series

The following CGEB-invited speakers gave seminars in the Department of Biochemistry & Molecular Biology, Department of Biology, or Faculty of Computer Science at Dalhousie over the past year. These were either sponsored (or co-sponsored) by funding from the Tula Foundation.

- (upcoming) Dr. David Liberles, University of Wyoming. "*A modeling framework for protein sequence evolution*"
- Dr. Christopher Hogue, National University of Singapore, September 2010. "*Toward a consistent evolutionary chronology of the ribosome*"
- Dr. David Roos, University of Pennsylvania, June 2010. "*Designing and mining (pathogen) genome databases*"
- Dr. Stephane Rombauts, University of Ghent, June 2010. "*Annotating a eukaryotic genome nowadays*"
- Dr. Bret Larget, University of Wisconsin, May 2010. "*A phylogeography of gorilla divergence*"
- Dr. Chris Lowe, University of Liverpool, April 2010. "*Landscape genetics for free-living microorganisms: genetic, genomic and ecophysiological approaches to understanding spatial diversity in marine protists*"
- Dr. David McMillen, University of Toronto, February 2010. "*Proof-of-principle disease prevention (in bacteria) through synthetic biology*"
- Dr. Graham Bell, McGill University, November 2009. "*The genomics of fitness in yeast*"

B. CGEB Meeting Sponsorship

CGEB made contributions to the following symposium through our *Seminar Series Fund* from the Tula Foundation:

- co-sponsored session speakers at the **CIFAR-IMB Workshop on Eukaryotic Genome Annotation** at Dalhousie University, Halifax, June 2010.
Speakers: David Roos (University of Pennsylvania) and Stephane Rombauts (University of Ghent)

Total symposia contribution from CGEB Seminar Series Fund: \$3,646.72

C. CGEB Joint Lab Meetings

In 2009 we instituted a new schedule of CGEB joint lab meetings that include all ten CGEB labs' faculty members, undergraduate and graduate students, postdocs, and lab personnel. These meetings are generally held on the first Thursday of every month and continue on a year-round basis.

The purpose of these meetings is to introduce CGEB faculty and trainees to the ever-expanding group of trainees in the various labs. Faculty members and/or their students and postdocs provide an overview of their research program in a seminar-style format, with adequate time allotted for questions, constructive criticism and open discussion throughout.

D. CGEB and SEB Journal Club

The *CGEB Journal Club* (initiated in 2007) has expanded to include all CGEB labs' trainees and faculty members who are interested in participating. The Journal Club meets on bi-weekly Tuesdays throughout the year, with a hiatus from July-August. Participants select and present a current research paper for analysis and discussion. Many of our trainees have stated that this experience has been invaluable and an integral part of their overall training.

We also have another more specialized journal club in *Statistical Evolutionary Bioinformatics (SEB)* that meets bi-weekly to discuss the latest literature and/or current topics in computational biology and genomics.

XV. FUTURE DIRECTIONS

At our recent biannual CGEB members meeting, we discussed at length how we could promote even more interaction/collaboration and stimulate productivity within the CGEB group. From this discussion, three important initiatives were identified:

A. As a University-wide Centre with members from three different Faculties spread across campus, the PI's and the trainees have to make special efforts to interact regularly to form collaboration. So far this interaction occurs mostly in the context of the CGEB seminar series and monthly group meetings. We believe that although extremely valuable, this does not provide enough dedicated time for the PIs and trainees in the various groups to learn about the various research interests of the groups. Borrowing from new CGEB member Claudio Slamovits' experiences as a trainee in the Tula sponsored Centre for Microbial Diversity and Evolution (CMDE) that holds an annual retreat, we propose to hold an annual (or biennial) 2-3 day CGEB retreat. This retreat will involve seminars/posters presented by CGEB members, Associates and trainees, as well as a handful of guest speakers (~70-80 individuals in total). Ideally this retreat would occur offsite at a remote resort hotel. This way all of those involved are more likely to attend all talks and participate in all events and the intensity of the interactions is thereby increased. By inviting several high profile speakers with similar research interests, connections could be forged with these outside groups and give CGEB a higher international profile. As part of this CGEB retreat, we will dedicate part of a day to 'brainstorm' about future collaborative research projects that CGEB is well-positioned to take on as a group (e.g. see **C.** below).

We will seek additional monies to support this annual (or biennial) activity as its projected cost (~\$30,000-40,000/year) extends well beyond the current CGEB budget. We feel this initiative is so important as to warrant using some of the 'New Opportunities' fund to leverage additional funds to make this retreat a reality.

B. As CGEB researchers are world-renowned for our work in the microbial biodiversity and phylogenomics, we think it would further enhance our profile and collaborative potential to sponsor an international Microbial Phylogenomics Meeting. An international meeting involving experts in microbial biodiversity, comparative genomics, phylogenetics and bioinformatics could be held at Dalhousie and some costs (e.g. invited speakers) could be offset by CGEB. A similar conference was held in Québec a few years ago and was a resounding success.

C. A critical component of data-driven research conducted by the CGEB group regards the stewardship of data and software. To achieve a lasting impact and foster collaborations, research data must be available through a common platform. Data itself is static over time, but the software infrastructure to manage and maintain the data requires dedicated work. This work is traditionally outside of the scope of research funding, yet must be performed by highly skilled individuals. High-impact research usually leverages freely accessible software platforms to create innovative and complex software infrastructure. The designers of this infrastructure are trainees who eventually leave to pursue new opportunities. As a result, the lasting value of databases and scientific applications depends on the ability to maintain the software as the platforms evolve. This issue is of strategic importance because it affects the entire research programs of CGEB

members rather than projects. It is further complicated by the fact that individual groups do not require a full-time commitment to stewardship and that funding agencies usually do not consider software maintenance within the scope of a funded project. The members of the CGEB have identified the requirements to support and maintain the software and data infrastructure. The best solution is to consolidate the activities by sharing the service of one of more highly qualified individuals whose task is to manage databases, update scientific software, and develop interfaces to data to make it accessible to researchers from a wide range of background. As discussed earlier in the 'Governance section' (page 7), a concrete 'deliverable' from this initiative would be a web-accessible CGEB microbial comparative genomics database and software platform based on Beiko's MOA system that will have more advanced functionality than any similar platform currently available worldwide.

APPENDIX A: GOVERNANCE

Proposed Governance of CGEB

The purpose of CGEB is to stimulate research and collaboration among its members. The streamlined management structure we propose is designed to ensure fair and efficient operation. The organizational structure of CGEB will consist of a Deans Committee, a Director, a Steering Committee, and the Membership. The terms of reference for these bodies are outlined below.

I. DEANS COMMITTEE

The Deans Committee will meet annually with the Director, providing advice on operation and direction of the Centre, and identifying emerging opportunities for CGEB to advance its research and training goals at Dalhousie University. The Deans committee will be headed by the 'Presiding Dean' (initially the Dean of the Faculty of Medicine) and, with the other Deans, will be ultimately responsible for the Centre management and budget. The Presiding Dean will also advise the Director on matters relating to the operation of the Centre as issues arise. The Deans Committee will appoint new Directors in consultation with the Steering Committee and the CGEB Membership. The Deans Committee will consist of: Vice-President, Research (or designate), ex-officio; Dean of Medicine (or designate), Presiding; Dean of Science (or designate); Dean of Computer Science (or designate).

II. DIRECTOR

The Director is ultimately responsible for decisions relating to Centre business and will chair the Steering Committee and supervise a Centre Administrator who will coordinate CGEB group meetings and meetings of the Deans Committee and Steering Committee, as well as prepare and receive Centre-related correspondence and assist in accounts-keeping. The Director will report to the Presiding Dean on all matters relating to Centre operation and budget and will meet with the Deans Committee annually. The Director position will normally be held for a three-year term but this term may be renewed on the recommendation of the Deans Committee. When required, a new Director will be selected by the Deans Committee, based on advice received from the CGEB membership.

III. STEERING COMMITTEE

The Steering Committee will be made up of the Director and two other faculty members in the CGEB group. The Steering Committee will be responsible for advising the Director on matters such as the disbursement of Tula Foundation-supported postdoctoral and graduate student positions to CGEB faculty, as outlined in the main text of the proposal. Furthermore, the Director will consult with the Steering Committee on any decisions substantially affecting CGEB before a final decision is made by the Director. The term for membership on the Steering Committee is three years, but reappointment of an individual whose term has been completed is possible providing the Membership approves.

APPENDIX A: GOVERNANCE (continued)

IV. CGEB MEMBERSHIP

The CGEB Membership will have scheduled meetings bi-annually to discuss Centre business, and make/vote on recommendations to the Steering Committee and the Deans Committee regarding new CGEB members, nominations for Director, Steering Committee membership, and new opportunities for CGEB.

Proposed Composition of the Initial CGEB Management Structure

I. DEANS COMMITTEE

- Vice-President, Research (or designate), ex-officio
- Dean of Medicine (or designate), Presiding
- Dean of Science (or designate)
- Dean of Computer Science (or designate)

II. DIRECTOR

- Andrew Roger, Biochemistry and Molecular Biology

III. STEERING COMMITTEE

- Andrew Roger, Biochemistry and Molecular Biology
- Michael W. Gray, Biochemistry and Molecular Biology
- W. Ford Doolittle, Biochemistry and Molecular Biology

IV. CGEB MEMBERSHIP

- Andrew Roger, Biochemistry and Molecular Biology
- Michael Gray, Biochemistry and Molecular Biology
- Ford Doolittle, Biochemistry and Molecular Biology
- John Archibald, Biochemistry and Molecular Biology
- Alastair Simpson, Biology
- Joseph Bielawski, Biology, and Mathematics and Statistics
- Robert Beiko, Computer Science
- Christian Blouin, Computer Science, and Biochemistry and Molecular Biology
- Ed Susko, Mathematics and Statistics
- (new) Assistant Professor, Biochemistry and Molecular Biology (job search underway)

V. CENTRE ADMINISTRATOR

- Wanda Danilchuk, Biochemistry and Molecular Biology

APPENDIX A (continued)

Additions to CGEB Governance Procedures (2009)

A. Recruitment and vetting of potential Tula-sponsored Trainees.

Once a CGEB member has found a suitable candidate to fill the Tula-sponsored postdoc or Ph.D. student position, he/she will send the CV of the candidate to the Director and Steering committee who will adjudicate the application. If they find that the candidate is acceptable as a Tula-sponsored trainee, the member is informed and the CGEB Administrator (Wanda Danilchuk) liaises with Research Services, Financial Services and the Tula Foundation to set up the appropriate accounts in anticipation of the trainee's start date.

B. Disbursement of Trainee Travel/Meeting Funds.

The \$10,000 annual budget for CGEB trainee travel/attendance at meetings is divided equally amongst CGEB labs. If any monies are left over from the previous year, these monies get rolled over into the next year and added to the new \$10,000 allotment. The total is then divided by ten (total number of CGEB labs) which becomes that year's 'allocation' for each lab.

C. Seminar Series.

All CGEB members are invited to suggest (to the Director) potential seminar series speakers covering their areas of interest. If the Director approves the choice, the speakers are invited by the CGEB member who will serve as host for the speaker's visit. While the speaker is at Dalhousie, CGEB members (and their trainees) are invited to meet with the speaker and/or have lunch/dinner with them in order to promote interaction and discussion. An honorarium of \$500 is given to each seminar speaker. Where possible, CGEB will partner with a member's home department to co-sponsor high-profile speakers.

The membership has decided to extend the Seminar series fund to cover sponsorship (or co-sponsorship) of dedicated symposia at meetings held at Dalhousie in the area of interest of several of the CGEB membership. These symposia can be part of larger international meetings held in Halifax. However, if CGEB sponsors or co-sponsors a symposium, all CGEB members and trainees may attend without paying meeting fees (although usually many of them are already registered to attend the meeting). CGEB members are free to make proposals of CGEB sponsorship of these symposia at any time and the Director decides on the eligibility in consultation with the membership and steering committee. The CGEB seminar speakers invited, and symposia sponsored, over the past year are listed on p. 32.

D. The 'CGEB Associate' position.

Last year we created a 'CGEB Associate' category defined as faculty members at Dalhousie University whose research is aligned closely with CGEB goals and who may be active collaborators of CGEB members, in some cases co-supervising trainees. The CGEB Associate status does not entitle the faculty member to any financial contributions from the Centre, nor entitle them to Tula-sponsored trainee positions. However, Associates and their trainees will be invited to all Centre activities (i.e. joint lab meetings, sponsored talks, symposia, workshops, etc.), can suggest seminar speakers, and will be able to participate in co-supervision of Tula-sponsored trainees with a CGEB member to whom the trainee monies are awarded. We hope

this category makes it possible to expand the impact and interactions of the Centre without diluting its financial resources. To become a CGEB Associate, a CGEB member will ‘sponsor’ the candidate and circulate their CV to the membership and provide a brief justification as to why the individual deserves Associate status. The Director will decide on the appointment, in consultation with the membership and the Steering Committee.