



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

ANNUAL REPORT

October 1, 2010 – October 1, 2011

TABLE OF CONTENTS

I.	Executive Summary	1-3
II.	CGEB Membership	4
III.	CGEB Governance	5
IV.	Financial Status of the Centre	6-7
V.	Supervision of Research Trainees	8-9
VI.	Publications by CGEB Faculty and Trainees	10-16
VII.	Invited and Other Presentations by CGEB Members	17-19
VIII.	Invited and Other Presentations by CGEB Trainees	20-23
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 Joint Lab Meetings	36
	C. CGEB and SEB Journal Club	37
XV.	Future Directions	38-39
	<i>APPENDIX A: CGEB Governance Policy</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 2009-2010, CGEB expanded to include three affiliated ‘CGEB Associate’ investigators and their groups from the faculty of Science: two from Mathematics and Statistics [Hong Gu and Chris Field] and one from Biology [Bob Lee]. New to this year is our appointment of Tier I CRC-sponsored faculty member in Biology, Dr. Julie LaRoche as a CGEB Associate. The main focus of the Centre’s research is understanding how microbial genomes evolve and diversify, however the research programs collectively span computational biology, biological oceanography, 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 three 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 October 2010-October 2011 period, CGEB researchers have published a total of 76 manuscripts in peer-reviewed journals (pp. 10-16), including contributions in high impact journals such as *Science*, *Current Biology*, *Genome Research*, *Proceedings of the National Academy of Sciences USA*, and *PLoS Biology*. The international profile of CGEB researchers is further demonstrated by the 31 invited presentations given at international conferences or invited seminars over the past year (pp. 17-19). Furthermore, in the past year CGEB members have been awarded or continue to hold 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. 20-23). They have presented their results at international meetings (a total of 38 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 to 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 26 bioinformatics software tools. These are 'open-source' software tools (or suites of tools) for analysis of genomic data have been created, published and made available (or soon will be) 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 ~\$2,077,135 of external grant support over 2010-2011 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, and 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 nine 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 trainees collectively over the past year increased to 54 (pp. 8-9).

CGEB seminar series.

The CGEB Seminar Series is in full swing and has hosted nine well-attended seminars this year from internationally recognized researchers (see p. 36) with several more scheduled to occur in 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.

Ongoing CGEB activities and future directions.

So far, group CGEB activities have consisted of a bi-weekly journal club, 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. In brief, four initiatives have been proposed: (A) To hold an annual CGEB retreat offsite; (B) CGEB sponsorship of an international comparative genomics, microbial biodiversity and phylogenomics meeting; (C) Establishment of support for a full time database and software developer to build/maintain sophisticated microbial genomics databases, software repositories and comparative genome analysis tools for the CGEB collective; and (D) A new intra-CGEB collaboration-based postdoc sponsorship program.

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

II. CGEB MEMBERSHIP

Director:

- Andrew Roger, Professor, Biochemistry and Molecular Biology

Administrator:

- Wanda Danilchuk, 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

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
- Julie LaRoche, Professor, Biology

III. 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 four 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 was held on August 26, 2011. Over the past year we have had regular CGEB members meetings (latest October 6, 2011) and have discussed issues relating to Centre structure and activities. These are briefly summarized below:

A. CGEB Associates

Three new 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)

Newly appointed associate for 2011:

- Dr. Julie LaRoche (Professor, Biology)

B. Obstacles 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 it has been relatively easy to procure 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 system 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, Gray, Slamovits, Roger and Simpson). **The barrier to this effort however remains our ability to fund dedicated staff members whose job is to develop/maintain MOA and other eukaryotic genome annotation/visualization tools and populate them with data from the CGEB labs. We also lack solid funding for system administration** to administer the huge computational resources upon which these software tools depend (e.g. the 3 large computer clusters within CGEB). The lack of funding for dedicated ‘cyber-infrastructure’ staff and the dim prospects for obtaining it through granting agencies (who prefer to fund trainees) remains the biggest challenge for the future growth of CGEB research efforts in comparative genomics.

IV. FINANCIAL STATUS OF THE CENTRE

The Centre's administration (i.e. 50% administration support, seminar series fund, and student travel to meetings fund) 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 1-4 have been received (see Table B. below).

A. Tula Foundation Contributions

Fiscal Year (August 1-July 30)	Yearly installment
Year 1: 2007-2008	\$50,000 (received)
Year 2: 2008-2009	\$50,000 (received)
Year 3: 2009-2010	\$50,000 (received)
Year 4: 2010-2011	\$50,000 (received)
Year 5: 2011-2012	\$50,000 (received)
Year 6: 2012-2013	\$50,000
Year 7: 2013-2014	\$50,000
Year 8: 2014-2015	\$50,000
	Total: \$400,000

The *Seminar Series Fund* (\$15,000/year) enables us to invite world-renowned scientists to present seminars at Dalhousie University. The *Travel to Meetings Fund* (\$10,000/year) 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 committed over \$3,000,000 to 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 (received)
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 members’ 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). 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* which resulted in a net income of \$39,181.46. Second, Drs. Doolittle & Gray had been maintaining an account from a conference in 1981 that had generated 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 net revenue of \$106.01.

The current balance in this account is **\$39,796.24**.

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).

V. SUPERVISION OF RESEARCH TRAINEES

A total of **54** trainees were supervised in CGEB members' labs over the past year.

Note: CGEB trainees that are **currently Tula-funded** are highlighted.

CGEB Faculty Member	Name of Trainee
Andrew Roger Biochemistry and Molecular Biology	Matt Brown, Postdoctoral Fellow
	Eleni Gentekaki, Postdoctoral Fellow
(W. Ford Doolittle) Biochemistry and Molecular Biology	Laura Eme, Postdoctoral Fellow (co-supervised with Doolittle & funded by Doolittle's postdoc allocation)
	Anastasios Tsaousis, Postdoctoral Fellow
	Tommy Harding, PhD student (co-supervised with A. Simpson)
	Michelle Leger, PhD student
	Dan Gaston, PhD student
	Martin Kolisko, PhD student (co-supervised with A. Simpson)
	Liwen Zou, PhD student (co-supervised with E. Susko)
	Courtney Stairs, PhD student
	Javier Alfaro, MSc student
	Grant Stevens, Undergraduate technician
John Archibald Biochemistry and Molecular Biology	Eunsoo Kim, Postdoctoral fellow
	Goro Tanifuji, Postdoctoral Fellow (co-supervised with M. Gray & funded by Gray's postdoc allocation)
	Julia Hopkins, Postdoctoral Fellow (co-supervised with M. Gray based on new Tula collaborative postdoc allocation)
	Takuro Nakayama, Postdoctoral Fellow
	Shinichiro Maruyama, Visiting Researcher (JSPS, Japan)
	Bruce Curtis, PhD student
	Christa Moore, PhD student
	Robert Eveleigh, MSc student (co-supervised with R. Beiko)
	Tyler Mills, Honours student
	Rebecca Gibault, Honours student
	Katherine Richman, Honours student
Joseph Bielawski, Biology	Katherine Dunn, Postdoctoral Research Associate
	Mahdi Shafiei, Postdoctoral Fellow
	Conor Meehan, Postdoctoral Fellow (co-supervised with R. Beiko)
	Joey Mingrone, PhD student (co-supervised with E. Susko)
	Wei Chen, MSc student (co-supervised with H. Gu)
	Yuriy Khalak, Undergraduate student
Christian Blouin, Computer Science	Jose Sergio Hleap Lozano, PhD student
	Swati Shankar, PhD student
	Haibin Liu, PhD student

CGEB Faculty Member	Name of Trainee
Robert Beiko, Computer Science	Conor Meehan, Postdoctoral Fellow (<i>co-supervised with J. Bielawski</i>)
	Morgan Langille, Postdoctoral Fellow
	Dennis Wong, PhD student
	Christopher Whidden, PhD student
	Donovan Parks, PhD student
	Norman MacDonald, PhD student
	Robert Eveleigh, Master's student (<i>co-supervised with J. Archibald</i>)
	Mike Porter, Master's student
	Timothy Mankowski, Software developer
	Kathryn Dunphy, Undergraduate software developer
Michael Gray Biochemistry & Molecular Biology	Goro Tanifuji, Postdoctoral Fellow (<i>co-supervised with J. Archibald</i>)
	Ryan Gawryluk, PhD student; PDF in Roger lab as of Oct. 1 st
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>)
	Gordon Lax, Visiting student researcher
	Qianqian Zhang, Visiting Scholar (VSGS student)
Claudio Slamovits Biochemistry & Molecular Biology	Shehre-Banoo Malik, CIFAR Junior Fellow
	Gillian Gile, Postdoctoral Fellow
	Renny Lee, PhD student
	Hugo Lai, Honours student
	Sarah Dowler, Co-op student
Edward Susko Mathematics and Statistics	Huaichun Wang, Postdoctoral Fellow
	Liwen Zou, PhD student (<i>co-supervised with A. Roger</i>)
	Joey Mingrone, PhD student (<i>co-supervised with J. Bielawski</i>)
	Goldis Radajabalipour, MSc student (<i>co-supervised with C. Field</i>)

VI. PUBLICATIONS BY CGEB FACULTY AND TRAINEES

[includes published, in press, and accepted papers: total of 76 for past year]

Note: CGEB faculty are highlighted in bold and CGEB trainees are underlined.

Abad, M.G., Long, Y., Willcox, A., Gott, J.M., **Gray, M.W.** and Jackman, J.E. (2011) A role for tRNA^{His} guanylyltransferase (Thg1)-like proteins from *Dictyostelium discoideum* in mitochondrial 5'-tRNA editing. *RNA* 17: 613-623.

Archibald, J.M. (2011) Origin of eukaryotic cells: 40 years on. *Symbiosis, in press* (invited commemorative review).

Archibald, J.M. (2011) Plastid origins. In: *Organelle genetics: evolution of organelle genomes and gene expression*, C. Bullerwell (ed.), Springer-Verlag, *in press*.

Archibald, J.M. & Richards, T.A. (2010) Gene transfer: anything goes in plant mitochondria. *BMC Biology* 8:147.

Barberà, M.J., Ruiz-Trillo, I., Tufts, J.Y., Bery, A., Silberman, J.D. and **Roger AJ.** (2010) *Sawyeria marylandensis* (Heterolobosea) has a hydrogenosome with novel metabolic properties. *Eukaryotic Cell* 9: 1913-24.

Bay, R.A. and **Bielawski, J.P.** (2011) Recombination detection under evolutionary scenarios relevant to functional divergence. *Journal of Molecular Evolution* (*accepted for publication*).

Beiko, R.G. (2011) Telling the whole story in a 10,000-genome world. *Biology Direct* 6: 34.

Bowman, S., Hubert, S., Higgins, B., Stone, C., Kimball, J., Borza, T., Bussey, J.T., Simpson, G., Kozera, C., Curtis, B.A., Hall, J.R., Hori, T.S., Feng, C.Y., Rise, M., Booman, M., Gamperl, A.K., Trippel, E., Symonds, J., Johnson, S.C. and Rise, M.L. (2011) An integrated approach to gene discovery and marker development in Atlantic cod (*Gadus morhua*). *Marine Biotechnology* 13: 242-255.

Brown, M.W., Silberman J.D. and Spiegel F.W. (2011) “Slime molds” among the Tubulinea (Amoebozoa): Molecular systematics and taxonomy of *Copromyxa*. *Protist* 162: 277-287.

Brown, M.W., Silberman, J.D. and Spiegel, F.W. (2010) A morphologically simple species of *Acrasis* (Heterolobosea, Excavata), *Acrasis helenhemmesae* n. sp. *Journal of Eukaryotic Microbiology* 57: 346-353.

Chan, C.X., **Beiko, R.G.** and Ragan, M.A. (2011) Lateral transfer of genes and gene fragments in *Staphylococcus* extends beyond mobile elements. *Journal of Bacteriology* 193: 3964-3977.

Cocquyt, E., Gile, G.H., Leliaert, F., Verbruggen, H., Keeling, P.J. and De Clerck, O. (2010) Complex phylogenetic distribution of a non-canonical code in green algae. *BMC Evolutionary Biology* 10: e327.

Corradi, N. and **Slamovits, C.H.** (2011) The intriguing nature of microsporidian genomes *Briefings in Functional Genomics* (Epub ahead of print, 10.1093/bfpg/elq032].

Curtis, B.A. and **Archibald, J.M.** (2010) A spliceosomal intron of mitochondrial DNA origin. *Current Biology* 20: R919-920.

Curtis, B.A. and **Archibald, J.M.** (2010) Problems and progress in understanding the origins of mitochondria and plastids. In: *Symbioses and Stress*, J. Seckbach & M. Grube (Eds.), pp. 41-62, Springer-Verlag.

Doolittle, W.F., Lukes, J., **Archibald, J.M.**, Keeling, P.J. and **Gray, M.W.** (2011) Comment on “Does constructive neutral evolution play an important role in the origin of cellular complexity?” *Bioessays* 33: 427-429.

Flegontov, P., **Gray, M. W.**, Burger, G. and Lukeš, J. (2011) Gene fragmentation: a key to mitochondrial genome evolution in Euglenozoa? *Current Genetics* 57: 225-232.

Garvin, M.R., **Bielawski, J.P.** and Gharrett, A.J. (2011) Positive Darwinian Selection in the Piston That Powers Proton Pumps in Complex I of the Mitochondria of Pacific Salmon. *PLoS ONE* 6(9): e24127.

Gaston, D., **Susko, E.** and **Roger, A.J.** (2011) A phylogenetic mixture model for the identification of functionally divergent protein residues. *Bioinformatics* 27: 2655-2663.

Gentekaki, E. and Lynn, D. (2010) Evidence for cryptic speciation in *Carchesium polypinum* Linnaeus 1758 (Ciliophora: Peritrichia) inferred from mitochondrial, nuclear, and morphological markers. *Journal of Eukaryotic Microbiology* 57: 508-519.

Gile, G.H. and **Slamovits, C.H.** (2011) Phylogenetic position of *Lophomonas striata* Bütschli (Parabasalida) from the hindgut of the cockroach *Periplaneta americana*. *Protist*, in press.

Gile, G.H., James, E.R., Scheffrahn, R.H., Carpenter, K.J., and Keeling, P.J. (2011) Molecular and morphological analysis of the Calonymphidae with a description of *Calonympha chia* sp. nov., *Snyderella kirbyi* sp. nov., *Snyderella swezyae* sp. nov., and *Snyderella yamini* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, in press.

Gray, M.W. (2011) The incredible shrinking organelle. *EMBO Reports* 12: 873.

Gray, M.W. and **Archibald, J.M.** (2011) Origins of mitochondria and plastids. In *Advances in photosynthesis and respiration (Genomics of Chloroplasts and Mitochondria)*, R. Bock and V. Knoop (Eds.), Springer, Dordrecht, in press.

Gray, M.W. and Beyer, A.L (2011) Ribonucleic acid (RNA). In *McGraw Hill Encyclopedia of Science & Technology, 11th Edition*, in press.

Gray, M.W., Lukes, J., **Archibald, J.M.**, Keeling, P.J. and **Doolittle, W.F.** (2010) Cell biology. Irremediable complexity? *Science* 330: 920-921.

Gu, H., Dunn, K.A. and **Bielawski, J.P.** (2011) Codon evolution: Mechanisms and models In: *Likelihood Based Clustering (LiBaC) for Codon Models* (Chapter 5), Gina M. Cannarozzi & Adrian Schneider (Eds.), Oxford University Press.

Halary, S.* , Malik, S.-B.*, Lildhar, L., **Slamovits, C.H.**, Hijri, M. and Corradi, N. (2011) Conserved meiotic machinery in *Glomus* spp., a putatively ancient asexual fungal lineage. *Genome Biology and Evolution* [Epub ahead of print, doi:10.1093/gbe/evr089]

Hampl, V., Stairs, C.W. and **Roger A.J.** (2011) The tangled past of eukaryotic enzymes involved in anaerobic metabolism. *Mobile Genetic Elements* 3: 71-74.

Harding, T., Jungblut, A.D., Lovejoy, C. and Vincent, W.F. (2011) Microbes in High Arctic snow and implications for the cold biosphere. *Applied and Environmental Microbiology* 77: 3234-3243.

Heiss, A.A., Walker, G. and **Simpson, A.G.** (2011) The ultrastructure of *Ancyromonas*, a eukaryote without supergroup affinities. *Protist* 162: 373-393.

Holloway, C. and **Beiko, R.G.** Assembling networks of microbial genomes using linear programming. *BMC Evolutionary Biology* 10: 360.

Kamikawa, R., Inagaki, Y., Tokoro, M., **Roger, A.J.** and Hashimoto, T. (2011) Split introns in the genome of *Giardia intestinalis* are excised by spliceosome-mediated trans-splicing. *Current Biology* 21: 311-315.

Kerney, R., Kim, E., Hangarter, R., Heiss, A.A., Bishop, C.D. and Hall, B.K. (2011) Intracellular invasion of green algae in a salamander host. *Proceedings of the National Academy of Sciences USA* 108: 6497–6502.

Kim, E.*, Harrison, J.W.* , Sudek, S.* , Jones, M.D.M., Wilcox, H.M., Richards, T.A., Worden, A.Z. and **Archibald, J.M.** (2011) Newly identified and diverse plastid-bearing branch on the eukaryotic tree of life. *Proceedings of the National Academy of Sciences USA* 108: 1496–1500.

Koenig, J.E., Bourne, D.G., Curtis, B., Dlutek, M., Stokes, H.W., **Doolittle, W.F.** and Boucher, Y. (2011) Coral-mucus-associated *Vibrio* integrons in the Great Barrier Reef: genomic hotspots for environmental adaptation. *ISME Journal* 5: 962-972.

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.** (2010) A wide diversity of previously undetected free-living relatives of diplomonads isolated from marine/saline habitats. *Environmental Microbiology* 12: 2700-2710.

Langelan, D.N., Wieczorek, M., **Blouin, C.** and Rainey, J.K. (2010) Improved helix and kink characterization in membrane proteins allows evaluation of kink sequence predictors. *Journal of Chemical Information and Modeling* 50: 2213-2220.

Liu, H., Keselj, V. and **Blouin, C.** (2011) Biological event extraction using subgraph matching. *Computational Intelligence, in press.*

Long, S., Changmai, P., Tsaousis, A.D., Skalický, T., Verner, Z., Wen, Y.Z., **Roger, A.J.** and Lukeš, J. (2011) Stage-specific requirement for Isa1 and Isa2 proteins in the mitochondrion of *Trypanosoma brucei* and heterologous rescue by human and *Blastocystis* orthologues. *Molecular Microbiology* 81: 1403-1418.

Lovejoy, T.E. et al. (13 co-authors, incl. **W.F. Doolittle**) (2010) *Canadian Taxonomy: Exploring Biodiversity, Creating Opportunity*. Canadian Council of Academies Press, Government of Canada.

Lowe, C.D., Keeling, P.J., Martin, L.E., **Slamovits, C.H.**, Watts, P.C. and Montagnes, D.J.S., (2011) Who is *Oxyrrhis marina*? Morphological and phylogenetic studies on an unusual dinoflagellate. *Journal of Plankton Research* (Epub ahead of print, doi:10.1093/plankt/fbq110)

Lukes, J., **Archibald, J.M.**, Keeling, P.J., **Doolittle, W.F.** and **Gray, M.W.** (2011) How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life* 63: 528-537.

Malik S.-B., Brochu, C.D., Bilic, I., Yuan, J., Hess, M., Logsdon, J.M. Jr. and Carlton, J.M. (2011) Phylogeny of parasitic Parabasalia and free-living relatives inferred from conventional markers vs. *Rpb1*, a single-copy gene. *PLOS – One*, 6(6): e20774.

Maruyama, S., Suzuki, T., Weber, A.P.M., **Archibald, J. M.** and Nozaki, H. (2011) Eukaryote-to-eukaryote gene transfer gives rise to genome mosaicism in euglenids. *BMC Evolutionary Biology* 11: 105.

Maruyama, S. and **Archibald, J.M.** (2011) Endosymbiosis, gene transfer and algal cell evolution. In *Advances in Algal Cell Biology*, K. Heimann & C. Katsaros (Eds.), *in press.*

McInerney, J.O., Martin, W.F., Koonin, E.V., Allen, J.F., Galperin, M.Y., Lane, N., **Archibald, J.M.** and Embley, T.M. (2011) Planctomycetes and eukaryotes: a case of analogy not homology. *BioEssays*. Doi 10.1002/bies.201100045.

Meehan, C.J. and **Beiko, R.G.** (2011) Habitat-directed lateral gene transfer of nickel transport proteins within the human gut microbiome. *Pacific Symposium on Biocomputing, accepted.*

Mora, C., Tittensor, D.P., Adl, S. and **Simpson A.G.B.** and Worm, B. (2011) How many species are there on Earth and in the Ocean? *PLOS Biology* 9: e1001127.

Park, J.S. and **Simpson A.G.B.** (2011) Characterization of *Pharyngomonas kirbyi* (= “*Macropharyngomonas halophila*” nomen nudum), a very deep-branching, obligately halophilic heterolobosean flagellate. *Protist* [available online 30 June 2011].

- 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* 57: 520-528.
- Parks, D.H.*, MacDonald, N.J.* and **Beiko, R.G.** (2011) Classifying short genomic fragments from novel lineages using composition and homology. *BMC Bioinformatics* 12: 328.
- Richards, T.A. & **Archibald J.M.** (2011) Cell evolution: gene transfer agents and the origin of mitochondria. *Current Biology* 21: R112-114.
- Saldarriaga, J., Gile, G.H., James, E.R., Horák, A., Scheffrahn, R.H., and Keeling, P.J. (2011) Morphology and molecular phylogeny of *Pseudotriconympha hertwigi* and *Pseudotriconympha paulistana* (Triconymphaea, Parabasalia) from neotropical rhinotermitids. *Journal of Eukaryotic Microbiology*, in press.
- Sangaralingam, A., **Susko, E.**, Bryant, D. and Spencer, M. (2010) On the artefactual parasitic eubacteria clain in conditioned logdet phylogenies: heterotachy and ortholog identification artefacts as explanations. *BMC Evolutionary Biology* 10: 343.
- Schnare, M.N. and **Gray, M.W.** (2011) Complete modification maps for the small and large subunit rRNAs of *Euglena gracilis*: functional and evolutionary implications of contrasting patterns between the two rRNA components. *Journal of Molecular Biology*, in press (accepted 16/08/11)
- Slamovits, C.H.** and Keeling, P.J., (2011) Contributions of *Oxyrrhis marina* to molecular biology, genomics and organelle evolution of dinoflagellates. *Journal of Plankton Research* (Epub ahead of print, doi: 10.1093/plankt/fbq153)
- Slamovits, C.H.**, Okamoto, N., Burri, L. James, E. and Keeling, P.J. (2011) A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nature Communications* 2:183.
- Spencer, D.F. and **Gray, M.W.** (2011) Ribosomal RNA genes in *Euglena gracilis* mitochondrial DNA: fragmented genes in a fragmented genome. *Molecular Genetics and Genomics* 285: 19-31.
- Spiegel, F.W., Shadwick, L., Brown, M.W., Nderitu, G., Aguliar, M., and Shadwick, J.D.L. (2012). Protostelida. In Margulis et al. (Eds.), *Handbook of Protoctista*, 2nd Edition, Jones and Bartlett Publishers, Boston – in press.
- Stairs, C.W., **Roger, A.J.**, and Hampl, V. (2011) Eukaryotic pyruvate formate lyase and its activating enzyme were acquired laterally from a firmicute. *Molecular Biology and Evolution* 28: 2087-2099.
- Stokesbury, M.J.W., Neilson, J.D., **Susko, E.** and Cooke, S.J. (2011) Estimating mortality of Atlantic bluefin tuna in an experimental recreational catch-and-release fishery. *Biological Conservation* 144: 2684-2691.

Susko, E. (2011) Improved least squares topology testing and estimation. *Systematic Biology* 60: 668-675.

Susko, E. (2011) Large sample approximations of probabilities of correct evolutionary tree estimation and biases of maximum likelihood estimation. *Statistical Applications in Genetics and Molecular Biology* 10(1): Article 10.

Tanifuji, G. (2011) Reductive genome evolution in nucleomorphs. *Japanese Journal of Protozoology – in press.*

Tanifuji, G., Kim, E., Onodera, N.T.*, Gibeault, R., Dlutek, M., Cawthorn, R.J., Fiala, I., Lukeš, J., Greenwood, S.J. and **Archibald, J.M.** (2011) Genomic characterization of *Neoparamoeba pemaquidensis* (Amoebozoa) and its kinetoplastid endosymbiont. *Eukaryotic Cell* 10: 1143–1146.

Tanifuji, G., Onodera, N.T., Wheeler, T.J., Dlutek, M., Donaher, N. and **Archibald, J.M.** (2011) Complete nucleomorph genome sequence of the non-photosynthetic alga *Cryptomonas paramecium* reveals a core nucleomorph gene set. *Genome Biology and Evolution* 3: 44-54.

Tong, J., Dolezal, P., Selkrig, J., Crawford, S., **Simpson, A.G.B.**, Noinaj, N., Buchanan, S.K., Gabriel, K. and Lithgow, T. (2011) Ancestral and derived protein import pathways in the mitochondrion of *Reclinomonas americana*. *Molecular Biology and Evolution*, **28**: 1581-1591.

Torruella, G., Derelle, R., Paps, J., Lang, B.F., **Roger, A.J.**, Shalchian-Tabrizi, K. and Ruiz-Trillo, I. (2011) Phylogenetic relationships within the Opisthokonta based on phylogenomic analyses of conserved single copy protein domains. *Molecular Biology and Evolution* [July 28 - Epub ahead of print, PMID: 21771718]

Tsaousis, A.D., Gaston, D., Stechmann, A., Walker, P.B., Lithgow, T. and **Roger, A.J.** (2011) A functional Tom70 in the human parasite *Blastocystis* sp.: Implications for the evolution of the mitochondrial import apparatus. *Molecular Biology and Evolution* 28: 781-791.

Tsaousis, A.D., Leger, M.L., Stairs, C.W., and **Roger, A.J.** (2011) The biochemical adaptations of mitochondrion-related organelles of parasitic and free-living microbial eukaryotes to low oxygen environments. In *Cellular Origins, Life in Extreme Habitats and Astrobiology Series, Vol. 21: Anoxia: Paleontological Strategies and Evidence for Eukaryote Survival*, A.V. Altenbach, J.M. Bernhard & J. Seckbach (Eds.), Springer – *in press.*

Veillette, J., Lovejoy, C., Potvin, M., Harding, T., Jungblut, A.D., Antoniadis, D., Chénard, C., Suttle, C.A. and Vincent, W.F. (2011) Milne Fiord epishelf lake: a coastal Arctic ecosystem vulnerable to climate change. *Ecoscience*: 18: *in press.*

Wang, H.-C., **Susko, E.** and **Roger, A.J.** (2011) Fast statistical tests for detecting heterotachy in protein evolution. *Molecular Biology and Evolution* 28: 2305-2315.

Wu, J. and **Susko, E.** (2011) A test for heterotachy using multiple pairs of sequences. *Molecular Biology and Evolution* 28: 1661-1673.

Zhaxybayeva, O. and **Doolittle, W.F.** (2011). Lateral gene transfer. *Current Biology* 21: R242-R246.

Zou, L., **Susko, E.**, Field, C. and **Roger, A.J.** (2011) The parameters of the Barry and Hartigan General Markov Models are statistically nonidentifiable. *Systematic Biology* [Apr. 20 - Epub ahead of print; PMID 21471307]

VII. INVITED AND OTHER PRESENTATIONS BY CGEB MEMBERS (total of 31)

Archibald, J.M. *Endosymbiosis and genome mosaicism in eukaryotic microorganisms.* Symposium entitled *Molecular Evolution in the Genomic Era*, Rome 3 University, Rome, Italy, September 2011 (invited talk)

Archibald, J.M. *One plus one equals one: secondary endosymbiosis and genome mosaicism in microbial eukaryotes.* 19th Annual Meeting of the Society for Molecular Biology and Evolution. Kyoto, Japan, July 2011 (invited talk)

Archibald, J.M. *Genome mosaicism in microbial eukaryotes.* Evolutionary Genomics Workshop, Institute for Pure & Applied Mathematics, UCLA, CA, November 2010 (invited talk)

Archibald, J.M. *Endosymbiosis and genome mosaicism in microbial eukaryotes.* University of New Brunswick, Department of Biology Seminar Series, February 2011 (invited seminar)

Archibald, J.M. (with B.A. Curtis, G. Tanifuji, **M.W. Gray**, P.J. Keeling, G.I. McFadden, C.E. Lane, J. Schmutz, A. Kuo and I. Grigoriev) *Nuclear genomes of Cryptophyte and Chlorarachniophyte.* Plant and Animal Genome XIX Conference, San Diego, CA, January, 2011 (invited talk)

Archibald, J.M. *Endosymbiosis and genome mosaicism in microbial eukaryotes.* Memorial Symposium for the 26th International Prize for Biology, Tsukuba, Japan, December 2010 (invited lecture)

Beiko, R.G. *GenGIS: A 3D geospatial environment for the analysis of genetic data.* iPlant GIS Meeting, Arizona, February 2011 (invited talk)

Beiko, R.G. *Bioinformatics of environmental samples: Who is there, what are they doing, and how will they respond?* Application of Next Generation DNA Sequencing in Environmental Risk Assessment and Monitoring: Future Challenges Meeting, Milan, Italy, May 2011 (invited talk)

Beiko, R.G. *Questions at the interface of ecology, evolution, space and time.* Annual Meeting of the Society for Molecular Biology and Evolution, Kyoto, Japan, July 2011 (talk)

Beiko, R.G. *Phylogenomics and lateral gene transfer in a 10,000-genome world.* Annual Meeting of the Society for Molecular Biology and Evolution, Kyoto, Japan, July 2011 (poster presentation)

Bielawski, J.P. *Workshop on Molecular Evolution – 1a.* Codon substitution models and phylogenetic analysis of protein coding genes. *1b.* Tutorial on use of PAML package of computer programs to analyze gene sequences for presence of sites evolving under adaptive evolution, Woods Hole, MA, July 2011 (invited lectures)

Bielawski, J.P. *Modeling and mapping microbial diversity and function with marker genes, genomes and metagenomes.* Canadian Microbiome Initiative Inaugural Team Meeting

Workshop, Vancouver, BC, March 2011 (invited talk)

Bielawski, J.P. *Searching for functional divergence in genes, genomes and metagenomes.* Modelling and Mapping Human Microbiome Diversity and Function, Halifax, May 2011 (invited talk)

Blouin, C. (with Haibin Liu and Vlado Keselj) *Biological event extraction using subgraph matching.* Fourth International Symposium on Semantic Mining in Biomedicine, Hinxton, Cambridgeshire, U.K., October 2010 (invited presentation)

Doolittle, W.F. Is there a universal Tree of Life? Department of Biology Seminar Series, *University of New Brunswick*, Fredericton, NB, September 2011 (invited seminar)

Doolittle, W.F. *Microbial Diversity Lectures*, Marine Biological Laboratory, Woods Hole, July 2011 (invited lectures)

Doolittle, W.F. CIHR Microbiomics Grant Workshop – Introduction and Overview, Dalhousie University, June 2011 (co-organizer and presenter)

Gile, G. and **Slamovits, C.H.** Phylogenetic position of *Lophomonas striata* Bütschli (Parabasalia) from the hindgut of the cockroach *Periplaneta americana*. European Congress of Protistology, Berlin, Germany, July 2011 (talk)

Gray, M.W. *The evolutionary origin of RNA editing.* Gordon Research Conference on RNA Editing: Editing and Modification of RNA and DNA, Galveston, TX, January 9-14, 2011 (invited talk)

Gu, H., Kenney, T., **Bielawski, J.** and Dunn, K. *Estimation of Darwinian positive selection using generalized codon-based models.* 39th Annual Meeting of the Statistical Society of Canada, Acadia University, Wolfville, N.S., June 12-15, 2011 (invited talk)

MacDonald, N., Parks, D. and **Beiko, R.G.** *RITA: Rapid Identification of high-confidence Taxonomic Assignments for metagenomic data.* International Human Microbiome Consortium Meeting, Vancouver, March 2011 (poster presentation)

Park, J.S., Grimm, K., Zhang, Q.Q., Harding, T., Brown, M.W., **Simpson, A.G.** *The deeper-level phylogeny and evolution of Heterolobosea: a major group of protists.* Joint meeting of the Phycological Society of America and the International Society of Protistologists, Seattle, WA, July 2011 (poster presentation)

Parks, D. and **Beiko, R.G.** *Using GenGIS to explore and test microbiome diversity at different spatial scales.* International Human Microbiome Consortium Meeting, Vancouver, March 2011 (poster presentation)

Roger, A.J. *Diversity in mitochondrion-related organelles in anaerobic protists; investigating evolutionary patterns and processes.* IUMS 2011, Sapporo, Japan, September 6-10, 2011 (invited talk)

Roger, A.J. *Diversity in mitochondrion-related organelles in anaerobic protists; investigating evolutionary patterns and processes.* National Institute of Infectious Disease, Tokyo, Japan, September 12, 2011 (invited talk)

Roger, A.J. *The evolution of protists and their organelles: new insights from the frontiers of genomics.* Joint Meeting of the Phycological Society of America and International Society of Protistologists, Seattle, WA, July 12-16, 2011 (invited Seymour H. Hunter Prize Lecture)

Roger, A.J. *Eukaryotic origins and the evolution of mitochondria.* Vienna Biocenter PhD Symposium on Origin of Life, Vienna, Austria, November 18-19, 2010 (invited talk)

Simpson A.G.B. *Eukaryote evolution; a story told by free-living protozoa.* 1st Asian Conference on Protistology, and 8th Asian Conference on Ciliate Biology, Jeju Island, Korea, September 2011(plenary lecture)

Susko, E. *Properties of Bayesian posteriors and bootstrap support in phylogenetic inference.* Centre de Recherches de Mathématiques, Université de Montréal, September 2011 (invited talk)

Susko, E. *Statistical Issues in Molecular Evolution: Measures of Uncertainty.* CRM-SSC Medal Address, Annual Meeting of the Statistical Society of Canada, Wolfville, Nova Scotia, June 2011 (invited talk)

Other Contributions:

- John Archibald was the Lead Organizer for *Guillardia theta* and *Bigelowiella natans* International Genome Jamboree, DOE Joint Genome Institute, Walnut Creek, CA, September 7-9, 2011.
- Robert Beiko and Donovan Parks co-organized a workshop entitled "Assessing and visualizing the geographic and temporal structure of biodiversity" at the *Annual Meeting of the Society for Molecular Biology and Evolution*, Kyoto, Japan, July 2011. Beiko also gave a presentation as part of this workshop – see above.
- Robert Beiko, Ford Doolittle, and Joe Bielawski co-organized the *CIHR Microbiomics Grant Workshop* at Dalhousie University, Computer Science Building, June 8-9, 2011. This included 25 participants from several CIHR-funded microbiome projects and several guest speakers.
- Ford Doolittle, Robert Beiko and Joe Bielawski co-organized a workshop entitled “*Metagenomics and microbiomics: concepts and methods*” at Dalhousie University, September 22, 2011. Participants included CGEB faculty and trainees, with invited seminar speaker Dr. Rob Knight, Memorial University.
- Roger co-convened a symposium on “*The evolution of parasitic protozoa*” at the International Union of Microbiological Societies Meeting in Sapporo, Japan, September, 2011.

VIII. INVITED AND OTHER PRESENTATIONS BY CGEB TRAINEES

This section includes posters, talks, and seminar presentations (**total of 38**) by CGEB trainees at universities, national and international meetings, conferences and workshops. Travel and attendance in some cases was sponsored in part by CGEB (Tula) funds.

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

Alfaro, J.* *Effects of protein structural constraints on models of protein evolution.* Biochemistry departmental seminar, May 12, 2011 (talk)

Brown, M.W.*, Silberman, J.D. and Spiegel, F.W. *A contemporary evaluation of the acrasids (Acrasidae, Heterolobosea, Excavata).* VI European Congress of Protistology, Freie Universität, Berlin, Germany, July 2011 (poster presentation)
***awarded Best Poster Presentation Prize for this presentation**

Brown, M.W.*, Silberman, J.D., Spiegel, F.W., Kolisko, M. and Roger, A.J. *Evolutionary history of aggregative multicellularity, insights from phylogenomics of Guttulinopsis.* VI European Congress of Protistology, Berlin, Germany July 2011 (talk)

Brown, M.W.* *New views on multicellularity from slime molds.* Biology Department seminar series, Dalhousie University, October 14, 2010 (invited seminar)

Curtis, B.E.* and Archibald, J.M. *Endosymbiotic gene transfer, endosymbiotic gene replacement and genome / proteome mosaicism. Guillardia theta and Bigelowiella natans* Genome Jamboree, DOE Joint Genome Institute, Walnut Creek, CA, September 2011 (talk)

Gawryluk, R.* *The molecular biology of mitochondria and chloroplasts.* Leopoldina Symposium on Molecular Genetics of Chloroplasts and Mitochondria, Berlin, Germany, October 2009 (poster presentation)

Gentekaki, E.* *Protistan phylogeography and population structure: Insights from the ciliated protist Carchesium polypinum. Brief introduction to genomics.* Chinese Academy of Sciences in Guangzhou, Guangdong, China, August 2011 (invited seminar and series of training sessions)

Gentekaki, E.* *Extensive lateral gene transfer in Blastocystis sp. Nand II strain.* Joint PSA/ISOP Meeting, Seattle, WA, July 2011 (talk)

Gentekaki, E.* *Protistan phylogeography, insights from the peritrich ciliate Carchesium polypinum.* Biology Department seminar series, Dalhousie University, November 25, 2010 (invited seminar)

Gile, G.H.* *Progress in bioinformatic characterization of the PPC proteome. Guillardia theta and Bigelowiella natans* Genome Jamboree, DOE Joint Genome Institute, Walnut Creek, CA, September 7-9, 2011 (invited talk)

Gile, G.H.* *Phylogenetic position of Lophomonas: Implications for character evolution in Parabasalia*. European Congress of Protistology, Berlin, Germany, July 2011 (invited talk)

Harding, T.*, Brown, M.W., Park, J.S., Roger, A.J. and Simpson, A.G.B. *The origin of the Heterolobosea: Insights from the amoeboid stage of Pharyngomonas*. VI European Congress of Protistology, Berlin, Germany, July 2011 (poster presentation)

Hleap, J.S.* and Blouin, C. *Defining evolutionary modules in protein structures*. 3DSIG 2011 - Structural Bioinformatics and Computational Biophysics Meeting (an ISMB satellite meeting), Vienna, Austria, July 15-16, 2011 (poster presentation)

Hopkins, J.F.* (with D.F. Spencer, M.W. Gray and J.M. Archibald) *Proteomics reveals complex evolution of plastid- and nucleomorph-targeted proteins in the chlorarachniophyte Bigelowiella natans*. *Guillardia theta* and *Bigelowiella natans* Genome Jamboree, DOE Joint Genome Institute, Walnut Creek, CA, September 7-9, 2011 (talk)

Kim, E.* *Protists: Diversity, symbiosis, and evolution*. American Museum of Natural History, NY, May 2011 (invited talk)

Kim, E.* and Archibald, J.M. *RNA-Seq Data*. *Guillardia theta* and *Bigelowiella natans* Genome Jamboree, DOE Joint Genome Institute, Walnut Creek, CA, September 2011 (talk)

Langille, M.* *Predicting function from taxonomy*. CIHR Microbiome Grant Workshop, Dalhousie University, June 2011 (invited talk)

Lee, R.* *Genome evolution in dinoflagellates*. Biochemistry departmental seminar, Dalhousie University, April 14, 2011 (seminar)

Lee, R.* *Workshop on analyzing next-generation sequencing data*. Michigan State University, East Lansing, MI, June 6-17, 2011 (invited participant)

Leger, M.M.*, Hug, L.A. and Roger, A.J. *A hydrogenosome in the free-living excavate *Andalucia incarcerationata*: Common themes in mitochondrial reduction in anaerobic eukaryotes*. VI European Congress of Protistology, Freie Universität, Berlin, Germany, July 2011 (invited talk)

Malik S.-B.*, Pightling A.W., Logsdon J.M. Jr., Slamovits C.H. and Archibald, J.M. *A model for meiosis and DNA repair in *Guillardia theta* and *Bigelowiella natans**. *Guillardia theta* and *Bigelowiella natans* Genome Jamboree, DOE Joint Genome Institute, Walnut Creek, CA, September 2011 (invited talk)

Malik S.-B.*, Halary S., Lildhar L., Hijri M., Corradi N., Pightling A.W., Logsdon J.M. Jr., Carlton J.M., Leander B.S. and Slamovits C.H. *New stories of sex in eukaryotes*. Center for Comparative Genomics & Evolutionary Bioinformatics, Dalhousie University, June 2, 2011 (seminar)

Maruyama, S.* and Archibald, J.M. *Sex-determining locus in G. theta. Guillardia theta and Bigeloviella natans* Genome Jamboree, DOE Joint Genome Institute, Walnut Creek, CA, September 2011 (talk)

Meehan, C.J.* *Microbiome metabolism mapping*. CIHR Microbiomics Grant Workshop, Dalhousie University, June 2011 (invited talk)

Moore, C.E.* and Archibald, J.M. *A small genome that 'thinks big': the complete nucleomorph genome of the cryptophyte alga Chroomonas*. Joint ISOP/PSA Meeting, Seattle, WA, July 2011 (talk)

***received the Eukaryotic Cell Young Investigator Award for this presentation**

*** received an ISOP Holtz-Connor Student Travel Award to attend this meeting**

Parks, D.H.* and Beiko, R.G. *GenGIS: A geospatial information system for genomic data*. Annual Meeting of the Society for Molecular Biology and Evolution, Kyoto, Japan, July 2011 (invited talk)

Parks, D.H.* and Beiko, R.G. *Quantifying beta-diversity over phylogenetic trees and networks*. Annual Meeting of the Society for Molecular Biology and Evolution, Kyoto, Japan, July 2011 (poster presentation)

Parks, D.H.* and Beiko, R.G. *Statistical analysis of metagenomic profiles*. Annual Meeting of the Society for Molecular Biology and Evolution, Kyoto, Japan, July 2011 (poster presentation)

Shafiei, M.*, Bielawski, J. and Gu, H. *Statistical models for network-structured data : Social networks and metabolic networks*. CIHR Microbiome Grant Workshop, Dalhousie University, June 8, 2011 (invited talk)

Stairs, C.W.*, Hampl, V. and Roger, A.J. *Acquisition of anaerobic metabolism: The radical case of pyruvate formate lyase*. Biochemistry departmental seminar series, Dalhousie University, May 5, 2011 (seminar)

Tanifuji, G.* *Examples for the genome study*. Special seminar in Department of Anatomy, Iwate Medical University, Iwate, Japan, August 2011 (invited departmental seminar)

Tanifuji, G.* and Archibald, J.M. *Comparative analysis of nucleomorph and nuclear genomes: investigation of the plastid and PPC proteomes. Guillardia theta and Bigeloviella natans* Genome Jamboree, DOE Joint Genome Institute, Walnut Creek, CA, September 2011 (talk)

Tanifuji, G.*, Onodera, N.T. and Hara, Y. *Nucleomorph genome diversity and its phylogenetic implications in cryptomonad algae* (talk). 35th Annual Meeting of the Japanese Society of Phycology, Toyama, Japan, March 26-28, 2011.

***received 14th Physiological Research Best Paper Award**

Tanifuji, G.*, Kim, E., Onodera, N.T., Gibeault, R., Dlutek, D., Cawthorn, R.J., Fiala, I., Lukeš, J., Greenwood, S.J. and Archibald, J.M. *Genomic characterization of Neoparamoeba pemaquidensis (Amoebozoa) and its kinetoplastid endosymbiont.* 19th Annual Meeting of the Society for Molecular Biology and Evolution, Kyoto, Japan, July 2011 (poster presentation)

Tsaousis, A.D.* *Reconstruction of the proteome of the mitochondrion-related organelle of Blastocystis sp.* EMBO Fellows Meeting, Heidelberg, Germany, June 16-19, 2011 (invited talk)

Tsaousis, A.D.* *The anaerobic biochemical adaptations of Blastocystis sp and its mitochondrion.* Department of Biology, University of Ottawa, Canada, January 6, 2011 (invited seminar)

Tsaousis, A.D.* *The biochemical adaptations of the mitochondrion-related organelle of Blastocystis sp.* EMBO Fellows Meeting, Salk Institute, La Jolla, San Diego, CA, November 5-7, 2010 (invited talk)

Whidden, C.* *Phylogenetic Distance Metrics.* Center for Comparative Genomics & Evolutionary Bioinformatics, Dalhousie University, March 3, 2011 (seminar)

IX. COLLABORATIVE INTERACTIONS BETWEEN CGEB MEMBERS

Collaborative interactions amongst CGEB labs continue to build. Evidence of this can be seen in the 8 publications that were co-authored by two or more CGEB members and their trainees in the last year (pp. 10-16). Below a short summary is provided of the collaborations that are ongoing within CGEB:

Archibald, Gray and Slamovits are working on the analysis and annotation of cryptophyte and chlorarachniophyte nuclear and organellar genome data coming from the Joint Genome Institute. Archibald and Gray are co-PIs on an NSERC Strategic Opportunities Grant to sponsor this work and co-supervise Tula-sponsored postdoc Dr. Goro Tanifuji. They also collaborate on the characterization of the nucleomorph proteome of cryptophytes and chlorarachniophytes, a task spear-headed by Tula-sponsored postdoc Dr. Julia Hopkins.

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 are working on the statistical estimation and comparison of model bias under tree topologies that yield artefacts and those that do not. They continue to collaborate on a new statistical framework for modeling complex microbial phenotypes as networks. They also work with Tula-sponsored postdoc Katherine Dunn in developing parametric modelling of heterogeneity in codon exchangeability rates, with investigation on model selection, model bias, and inference of sites subject to positive selection. As part of the human microbiome project, Bielawski, Gu, Dunn and postdoc Dr. Mahdi Shafiei are exploring Bayesian approaches to the modelling of the latent metabolic capacity of microbiomes.

Bielawski and Doolittle will hire a postdoc from the CIHR microbiome grant to apply the structured coalescent to the question of trans species gene pools in prokaryotes. This work has the potential to inform the philosophical connection between population genomics and the notion that prokaryotic species have indefinite boundaries.

Bielawski and Susko are collaborating on modelling covarion-like evolution at the codon level. They will also further develop codon models that have explicit population genetic parameters (e.g., population size and selection coefficients). Finally, they work together on the application of the bootstrap to the problem of site identification under codon models. As part of this work they are co-supervising the Ph.D. 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 are holders of a three-year CIHR Emerging Team Grant entitled “Modeling and mapping microbial diversity and function with marker genes, genomes and metagenomes”. This project has sponsored the hiring of three new postdocs that bring fresh expertise into the CGEB community.

Roger, Beiko and Archibald began a collaboration to understand the pattern and impact of lateral gene transfer in the anaerobic human parasite *Blastocystis*. Tula-sponsored postdoc Dr. Eleni Gentekaki leads this project.

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 the Excavata eukaryote super-group as well as other unplaced protist lineages by EST surveys and phylogenomic analysis. They co-supervise Ph.D. student Martin Kolisko who is embarking on an ambitious full-genome characterization of a 'missing-link' excavate organism *Carpediemonas membranifera*. Tula-sponsored postdoc Matt Brown works with Roger and Simpson on ultrastructure and phylogenomics of breviate amoebae.

Roger and Simpson have a collaboration to examine molecular evolution in very halophilic (salt-loving) protozoa, and to use transcriptomic surveys to shed light on their adaptation to extreme salinity. They co-supervise Ph.D. student Tommy Harding who is the main researcher working on this project.

Slamovits and Simpson are collaborating to culture and characterize new lineages of heterotrophic basal Alveolates (e.g. Colpodellids), to find suitable targets for further transcriptomic or genomic survey studies.

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. Tula-sponsored postdoc Dr. Huaichun Wang spearheads a new project on incorporating selection parameters into protein evolution models. Ph.D. student (Roger lab) Daniel Gaston continues to expand his FunDi framework for studying functional divergence and M.Sc. student Javier Alfaro (Roger lab) is working on structure-based phylogenetic models in collaboration with Susko and Blouin.

Susko, Roger and Field (Associate) have been developing maximally flexible general Markov models for sequence evolution, 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. This included work done both with Dr. Huaichun Wang and recently graduated Ph.D. student Dr. Liwen Zou.

X. CGEB DEVELOPED SOFTWARE

2010-2011:

Program name: *RITA* (Rapid Identification of Taxonomic Assignments)

Author: Robert Beiko

Description: A hybrid classification system that uses composition and homology information to assign metagenomic DNA fragments to their appropriate originating genomes. The Web version is available at <http://ratite.cs.dal.ca/rita>, and a standalone at <http://kiwi.cs.dal.ca/Software/RITA>.

Program name: *glsphyl, glsphylest, wlsnphyl, wlsphylest*

Author: Ed Susko

Description: Software for estimating trees and constructing confidence intervals for trees using least squares distance methods.

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

Program name: *Proteus*

Authors: Joe Bielawski and Joey Mingrone

Description: The code base has grown, now containing > 45,000 lines of code. The software has stationary and non-stationary codon models implemented under both a likelihood and, albeit crude, Bayesian framework. Covarion models are close to being completed. Anticipated first release within one year.

Program name: *Codon Optimal Likelihood Discoverer (COLD)*

Authors: Joe Bielawski, Kathy Dunn and Toby Kenney

Description: A program that calculates and maximizes the log-likelihood for complex user-defined codon models.

Website: <http://www.mscs.dal.ca/~tkenney/Cold/>

Program name: *lenergy*

Authors: Ed Susko, Javier Alfaro and Andrew Roger

Description: A maximum likelihood method that implements novel independent sites, structure-based, free-energy models of protein evolution.

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.

<http://rogerlab.biochemistryandmolecularbiology.dal.ca/fundi.php>

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

Authors: Donovan Parks, Mike Porter, Sylvia Churcher, S. Wang, Christian Blouin, J. Whalley, S. Brooks and Robert Beiko

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.
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.

<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.

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.

<http://www.mathstat.dal.ca/~hcwang/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.

<http://www.mathstat.dal.ca/~hcwang/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.

<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.

<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.

<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.

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

Program: *Radié*

Authors: J. Whalley, S. Brooks and Robert Beiko

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

<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.

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.

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.

<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.

<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.

<http://www.bielawski.info>

Program: *spr distance*

Authors: D. Hickey, Dehne, Rau-Chaplin and Christian Blouin

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

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

XI. CURRENT EXTERNAL RESEARCH GRANTS AND FUNDING

Collectively, CGEB faculty members received [~ \$2,077,135] from external research grants and other funding in 2010-11. Grants involving more than one CGEB faculty member as co-applicants were only counted once (i.e. for the primary P.I.) to obtain the above total.

John Archibald:

- 2011-2015 **Canadian Institutes of Health Research – Operating Grant:** *Endosymbiosis, parasitism, and genome evolution*: \$115,000/year
- 2011 **CIHR & NS/CIHR Regional Partnership Program Operating Grant —** *Endosymbiosis, parasitism, and genome evolution*: \$37,824 over 3 months
- 2009-2013 **NSERC Discovery Grant –** *Genome and proteome evolution in nucleomorph-containing algae*: \$34,000/year
- 2007-2012 **Tula Foundation,** CGEB Molecular Biology Research Fellowship (funding for postdoctoral fellowship plus research allowance): \$64,000/year
- 2007-2012 **Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity—**research allowance for program members: \$25,000/year

Robert Beiko:

- 2011-2014 **Genome Canada, Large-Scale Applied Research Projects:** *Biomonitoring 2.0: A high-throughput genomics approach for comprehensive biological assessment of environmental change* (co-applicant with M. Hajibabaei, Guelph, P.I.): ~\$3 million (not counted in above total)
- 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)
- 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 counted in above total)
- 2008-2013 **Tula Foundation –** funding for PhD student traineeship plus research allowance: \$27,500/year
- 2010-2012 **Tula Foundation –** funding for PhD student traineeship plus research allowance: \$27,500/year
- 2007-2012 **NSERC - Discovery Grant:** *New computational methods for metagenomics*: \$23,000/year
- 2007-2012 **Canada Research Chairs (CIHR) –** *Canada Research Chair (Tier II) in Bioinformatics*: \$100,000/year

Joseph Bielawski:

- 2011-2012 **ACEnet/Sun Microsystems Research Fellowship:** \$20,000
- 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)

- 2009-2013 **NSERC – Discovery Grant:** *Evolutionary analysis of protein diversification: model improvement, assessment and application to real data:* \$36,250/year
- 2007-2012 **Tula Foundation** – funding for CGEB postdoctoral fellowship plus research allowance: \$49,000/year

Christian Blouin:

- 2010-2015 **NSERC – Individual Discovery Grant:** *Exploring the landscape of phylogenies:* \$24,000/year
- 2009-2014 **CFI (Infrastructure Operating Funds):** *A high performance computational platform for bioinformatics:* \$12,000/year
- 2010-2015 **Tula Foundation** – CGEB funding for PhD student traineeships plus research allowance: \$27,500/year
- 2008-2013 **Tula Foundation** – CGEB funding for PhD student traineeship plus research allowance: \$27,500/year

W. Ford 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. Bielawski, R. Beiko, M. Ereshefsky)
- 2008-2010 **Tula Foundation,** CGEB Molecular Biology Research Fellowship (funding for postdoctoral fellowship plus research expenses): \$64,000/year for 2.5 years
- some funds still maintained in this account

Michael Gray:

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

Andrew Roger:

- 2011-2016 **Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant** – *Phylogenomic approaches to ancient relationships amongst eukaryotes:* 54,000/year (plus \$120,000 over 3 years as an Accelerator Supplement)
- 2011-2012 **Canada Foundation for Innovation/NSRIT/Dalhousie University Faculty of Medicine** – *A new laboratory for comparative genomics and evolutionary bioinformatics:* \$301,664 (renovations and equipment grant accompanying CRC; currently awaiting final approval by CFI, with matching funds from NSRIT)
- 2010-2017 **Canada Research Chairs – Tier I (CIHR) in Comparative Genomics and Evolutionary Bioinformatics:** \$200,000/year (includes 7-yr. salary award, 20% of which is research allowance)
- 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 allowance for program members: \$25,000/year

2007-2015 **Tula Foundation**, Centre for Comparative Genomics & Evolutionary Bioinformatics (CGEB) - administration, seminar series and student travel funds: \$50,000/year

2007-2012 **Tula Foundation**, CGEB Molecular Biology Research Fellowship (funding for postdoctoral fellowship plus research expenses): \$64,000/year

Alastair 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): \$64,000/year

2008-2013 **CIHR Operating Grant** (*co-applicant with A.J. Roger, P.I.:* \$143,876/year – *not counted in above total*)

2007-2011 **Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity** - research allowance for program members: \$25,000/year

Claudio Slamovits:

2011-2016 **Canada Foundation for Innovation – Leaders Opportunity Fund:** *A laboratory for microbial parasite genetics and evolution:* \$5,000/year (maintenance fund)

2010-2015 **NSERC Discovery Grant:** *Genomes of alveolate protists: structure, function and evolution:* \$27,000/year

2009-2014 **Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity**— research allowance for program members: \$25,000/year

Ed 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): \$49,000/year

XII. HONOURS/AWARDS/DISTINCTIONS

[includes those received from 2010-2011, 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 - renewed for 2012-2017
- appointment to Tenure Track (2011)

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:

- Member, Scholarly Integrity Policy Committee, Dalhousie University (2010-2011)
- 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
- President-Elect, International Society for Evolutionary Protistology (ISEP), 2010-2012
- Director, Centre for Comparative Genomics & Evolutionary Bioinformatics (CGEB), Dalhousie University (2008-2011; renewed for 2011-2014)
- Fellow, Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program (2007-2012)

Alastair Simpson:

- Fellow, Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program (2009-present)

Ed Susko:

- recipient of the CRM-SSC Prize (2011) from the Centre de recherches mathématiques (CRM) and the Statistical Society of Canada (SSC) in recognition of outstanding research accomplishments by a statistical scientist within the first 15 years after earning a doctorate

XIII. SERVICE TO THE SCIENTIFIC COMMUNITY (OUTSIDE DALHOUSIE)

A. Journal Editorial Boards (2010-2011)

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

In addition, 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 and International Service on Scientific Advisory Boards / Committees / Grant Review Panels (2010-2011)

- J. Archibald: Treasurer, **International Society for Molecular Biology and Evolution**
- J. Archibald: Peer Review Committee Member, **Canadian Institutes of Health Research – New Investigator Awards (Panel C) and Doctoral Research Award A (DRA) and Masters Awards (CGA)**
- J. Archibald: Organizing Committee Member, Marine Microbial Eukaryote Transcriptome Project (**Gordon and Betty Moore Foundation, and the Center for Genome Resources**)
- R. Beiko: Grant Reviewer for **NSERC, MITACS, and the Leenards Foundation**
- J. Bielawski: Steering Committee Member, **Marine Biological Laboratory Workshop on Molecular Evolution**
- J. Bielawski: Advisory Board Member, **European Workshop on Molecular Evolution**
- J. Bielawski: Referee for **NSERC and Swiss National Science Foundation (SNSF)**
- C. Blouin: Committee Member, **Atlantic Association for Research in Mathematical Science**
- W.F. Doolittle: Chair, Scientific Advisory Board, **Hydrocarbon Metagenomics Project – Genome Alberta: 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 and M.W. Gray: Advisory Committee Members, **Canadian Institute for Advanced Research – Integrated Microbial Biodiversity Program**
- A.J. Roger: Grant Reviewer for **National Science Foundation (USA) and Canadian Institutes of Health Research**
- A.J. Roger: President-Elect, **International Society for Evolutionary Protistology (ISEP)**
- A.G.B. Simpson: Executive Committee Member-at-Large, Program Chair, and Awards Committee Member for **International Society of Protistologists**
- A.G.B. Simpson: Grant Reviewer for **Czech Academy of Sciences, Leverhulme Trust (UK), NASA Exobiology Program and National Science Foundation (USA)**
- E. Susko: Grant Reviewer for **National Science Foundation (USA) and NSERC**
- E. Susko: Consultant for **Department of Fisheries and Oceans**

XIV. OTHER ACTIVITIES OF THE CENTRE

A. CGEB Seminar Series

The following invited speakers contributed to our seminar series at Dalhousie over the past year. These were either sponsored (or co-sponsored) by funding from the Tula Foundation.

- Dr. Rob Knight, University of Colorado: "*Spatial and temporal variability in the human microbiome*", September 22, 2011
- Dr. Andrew Lang, Memorial University: "*Gene transfer agents: capturing phage and putting them to work*", September 15, 2011
- Dr. Elisabeth Tillier, University of Toronto: "*Coevolution reveals a network of human proteins involved in ciliopathies and cancer*", August 9, 2011
- Dr. Yuji Inagaki, University of Tsukuba, Japan: "*Recent progress in placing novel protistan lineages in the tree of eukaryotes*", August 5, 2011
- Dr. Donald Baird, Environment Canada and UNB, and Dr. Mehrdad Hajibabaei, University of Guelph: "*Biomonitoring 2.0: A high-throughput genomics approach for comprehensive biological assessment of environmental change*", August 4, 2011
- Dr. Trevor Lawley, Wellcome Trust Sanger Institute, UK: "*Transmission and evolution of Clostridium difficile*", June 9, 2011
- Prof. Keith Willison, Institute of Cancer Research, UK: "*Analysis of the evolution of the eukaryotic CCT-actin folding system using structural and network biology data*", May 4, 2011
- Dr. Mary Beth Saffo, Marine Biological Laboratory, Woods Hole, MA: "*Mutualism from parasitic roots: a surprising three-kingdom symbiosis in the marine invertebrate family Molgulidae*", December 2010
- Dr. David Liberles, University of Wyoming: "*A modelling framework for protein sequence evolution*", November 8, 2010

B. 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.

C. 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 biweekly 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 biweekly to discuss the latest literature and/or current topics in computational biology, protein evolution and genomics.

XV. FUTURE DIRECTIONS

At our recent CGEB members meeting on Oct. 6th, we discussed at length how we could promote even more interaction/collaboration and stimulate productivity within CGEB:

A. Annual Retreat

This retreat would 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 bi-annual) activity as its projected cost (~\$30,000-40,000/year) extends well beyond the current CGEB Budget and will use some of the ‘New Opportunities’ fund to leverage additional funds.

B. A CGEB Database and Software Developer

A critical component of data-driven research conducted by the CGEB group regards the stewardship of data, databases and software. Unfortunately stable and long-term financial support for a highly skilled database and software developer is practically impossible to obtain through normal Canadian scientific granting councils. Instead many of the CGEB PIs maintain their databases and software tools ‘by the seat of their pants’ through the efforts of graduate students and postdocs. This compromises the research productivity of the trainees, and it does not provide long-term consistency or allow for optimal design, continued development and maintenance of key computational resources.

The best solution for CGEB is to consolidate our individual activities by sharing the service of one highly qualified individual whose task would be to manage databases, update scientific software, and develop interfaces to data to make it accessible to researchers from a wide range of backgrounds. As discussed earlier in the ‘Governance section’ (p. 5), 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. In addition, to this staff member, we have a desperate need for part-time support for system administration and information technology for the three high-performance computing clusters owned by the CGEB group. Monies to support these positions will be sought within the next year.

C. Collaborative Postdocs/PhD students Program

As we are past the halfway point in the five-year Centre renewal cycle, we have begun to think of the future of the most impactful of CGEB initiatives – our Tula supported postdoc/PhD program. In fact, the end of the funding period for some of the first postdoc/PhD student slots will be in 2012. There is no question how important this program has been in the career development of the trainees who have directly been sponsored, as well as the careers of the CGEB PIs. Therefore, we would like to continue the program in the future. The proposal we have settled on sees a slightly greater overall

number of postdoc/PhD student slots funded for five years (funded over an eight year period) as in the first round. However, to further foster CGEB integration, these trainee positions would be **specifically aimed at projects that represent collaborations between two or more CGEB Members or one Member and one or more Associates**. CGEB Director Roger will investigate logistics for the funding of this initiative within the next few months. Of all of the proposals for future directions, this one is by far the most important to the long-term vitality of CGEB.

D. International Phylogenomics Meeting

As CGEB researchers are world-renowned for our work in microbial biodiversity and phylogenomics, we think it would further enhance these fields and our 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. The idea is that perhaps two such international conferences could be convened in the next six years – every third year *in lieu* of the annual retreat. The monies for the annual retreat would instead underwrite the travel of invited speakers or other conference costs that are not recoverable by registration fees. A similar conference was held in Québec a few years ago and was a resounding success.

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 over the past year are listed on p. 35.

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.