BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Chip Stewart

eRA COMMONS USER NAME (credential, e.g., agency login): CHIP6286

POSITION TITLE: Associate Director, Scientific Projects

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
University of Illinois, Champaign Urbana, IL	B.S.	1981	Physics
Indiana University, Bloomington, IN	M.S.	1983	Physics
Indiana University, Bloomington, IN	Ph.D.	1988	High Energy Physics
Fermi National Accelerator Laboratory, Batavia, IL	Postdoc	1992	High Energy Physics
University of Michigan, Ann Arbor, MI	Postdoc	1995	High Energy Physics
Boston College, Chestnut Hill, MA	Postdoc	2008	Bioinformatics

A. Personal Statement

My role in the proposed research will be to advise and assist with data processing for somatic variant discovery from next generation sequencing data, and the identification of variants that play a significant role in tumorigenesis. Over the past decade I have developed software packages for structural variation discovery used in the 1000 Genomes project and for cancer genome analysis at the Broad Institute which have been used in many cancer projects and publications. In summary, I have a demonstrated record of successful and productive research projects in areas relevant to the proposed project

B. Positions and Honors

Positions and Employment

1995-2001 Research Scientist, Biologic Systems Corp., Mundelein, IL
2001-2005 Research Scientist, Massachusetts Institute of Technology Cambridge, MA
2008-2011 Assistant Research Professor, Boston College, Chestnut Hill, MA
2011-2018 Computational Biologist, The Broad Institute, Cambridge, MA
2018- Associate Director, Cancer Genome Analysis, The Broad Institute, Cambridge, MA

Honors

1986 Outstanding Graduate Student in Research, Indiana University Physics
2001 Outstanding Technical Contributions to Bio-logic Product Development
2015 Excellence Award in Mentorship/Teaching/Training, The Broad Institute

C. Contribution to Science

1. My early work in high-energy physics under the direction of Dr. Andzrej Zieminski at Indiana University quantified the onset of the quark parton model of proton-nucleus collisions at high energy. Our experiment, E557 at Fermilab, took advantage of the 800 GeV proton beam produced by the newly completely Tevatron at Fermilab to make the highest energy proton-nucleus collisions ever produced at that time. We surrounded the target collision point with an array of detectors to capture all particles produced in the collisions, including a large calorimeter and muon chambers designed, built, and tested at Indiana prior to being installed at Fermilab. Prior experiments observed showers of particles in all directors produced by proton-nucleus collisions at lower energies. The parton model, proposed by Richard Feyman and Rick Field in the early seventies, predicted that

"jets" of particles should be produced when the collision energy became sufficiently high such that quarks within protons behave as independent free quarks. Our experiment was the first to observe this behavior and was the basis of my first publication in the Physical Review in 1987. I happily continued my pursuit of particle physics as a post-doc at Fermilab where I made major contributions to the discovery of the top quark in 1995.

a. Gomez, R., Dauwe, L., Haggerty, H., Malamud, E., Nikolic, M., Hagopian, S., et al. (1987). Measurement of the nuclear enhancement in high-Et and jet event production. Physical Review D: Particles and Fields, 35(9), 2736–2746. PubMed PMID: 9957982
b. Stewart, C., Zieminski, A., Blessing, S., Crittenden, R., Draper, P., Dzierba, A., et al. (1990). Production of high-pt jets in hadron-nucleus collisions. Physical Review D: Particles and Fields, 42(5), 1385–1395. doi:10.1103/PhysRevD.42.1385. PubMed PMID: 10012978
c. Abachi, S., et al. (1995). Observation of the Top Quark. Physical Review Letters, 74(14), 2632–2637. doi:10.1103/PhysRevLett.74.2632. PubMed PMID: 10057979

2. As the discovery of the top quark approached I began to consider working in industry and found a position at Biologic Systems. At Biologic I developed software to interpret and display digital EEG information to assist clinicians in the diagnosis of sleep disorders, epilepsy, and as a diagnostic hearing test for newborn infants. I enjoyed my time Biologic, in particular learning from Dr. Ivan Pal, a neurologist with a keen interest in quantitative analysis of medical data, however I missed academic research so I joined the biophysics lab of James Weaver at MIT to develop transport models for ions across cell membranes (electroporation) and heat across bulk tissue. While at MIT I was attracted by the excitement of advances in genetics and the human genome project. In 2006 I took a post-doc position in the lab of Gabor Marth at Boston College. My primary focus was on the development of novel algorithms to detect Structural Variants (SV) in next generation sequencing data. At first my algorithms were developed and optimized using only simulated data because our lab did not have access to real sequencing data, but in 2008 we joined the 1000 Genomes Project and I became a founding member of the 1000G SV group under the leadership of Matt Hurles. Many institutions participated in the 1000G and contributed sequenced data and variant calls. I contributed SV calls from my algorithm "Spanner" which was consistently among the top algorithms in terms of SV sensitivity and false detection rate. I also work closely with two graduate students at Boston College, Michael Stromberg and Deniz Kural, to develop a split-read algorithm based on Michael's MOSAIK aligner software to detect mobile element insertions in 454 sequencing data. This algorithm was also among the best in terms of sensitivity and false detection rate. Spanner and Mosaik were included in the first 1000G publications and led to our own publication focused on population variation of mobile element insertions.

a. Stewart, D. A., Gowrishankar, T. R., Smith, K. C., & Weaver, J. C. (2005). Cylindrical cell membranes in uniform applied electric fields: validation of a transport lattice method. IEEE Transactions on Bio-Medical Engineering, 52(10), 1643–1653. doi:10.1109/TBME.2005.856030; PubMed PMID: 16235650
b. 1000 Genomes Project Consortium, Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., et al. (2010). A map of human genome variation from population-scale sequencing. Nature, 467(7319), 1061–1073. PubMed Central PMCID: PMC3042601; PubMed PMID: 20981092
c. Mills, R. E., Walter, K., Stewart, C., Handsaker, R. E., Chen, K., Alkan, C., et al. (2012). Mapping copy number variation by population-scale genome sequencing. Nature, 470(7332), 59–65. PubMed Central PMCID: PMC3077050; PubMed PMID: 21293372
Stewart, C., Kural, D., Stromberg, M. P., Walker, J. A., Konkel, M. K., Stütz, A. M., et al. (2011). A comprehensive map of mobile element insertion polymorphisms in humans. PLoS Genetics, 7(8), e1002236. PubMed Central PMCID: PMC3158055; PubMed PMID: 21876680

3. I joined the Cancer Genome Analysis group at the Broad Institute under the direction of Gad Getz in 2011. My first projects at Broad were focused on the analysis of pediatric cancers, Rhabdoid Tumors and Ewings Sarcoma. Pediatric tumors are usually associated with very low mutation rates, so I soon discovered that we needed new filtering schemes to reduce the level of artifacts to reduce the mutation false detection rate to a few percent. The need for high quality mutation calls led to the discovery of and subsequent filter for oxidative damage incurred in library prep resulting in a large number of G>T "OxoG" artifact mutations at low allele fraction. The oxidative damage was characteristic of exome data sequenced at Broad starting in late 2011. In 2012 the library prep protocol was modified to reduce oxidative damage, but occasionally we still find clear evidence of OxoG artifacts in sequenced data produced at other centers and our OxoG filter is able to reduce

the false detection rate to 1% while retaining nearly all true mutations. We published papers on Rhabdoid Tumors and Ewing Sarcoma as well as a paper describing our OxoG problem and solution. In 2012 I joined the TCGA analysis-working-group for papillary thyroid cancer as the “analysis coordinator”. This was a large sequencing project with a cohort of 500 patients. Once again, thyroid cancer is known as a low mutation rate tumor type and filtering strategies became an important theme. In the thyroid data we discovered a serious barcode swapping problem in library prep that was eventually resolved by introducing double bar-codes to maintain the identity of DNA fragments during library construction and sequencing, but in the mean time we needed a special bar-code post-process that identified instances of the bar-code swap in mutation calls and restored true mutations. This algorithm was also used for other affected projects such as the murine small cell lung carcinoma project. Since 2011 I’ve developed software tools for the CGA group which were used by analysts in most large-scale cancer sequencing projects at Broad. Currently my focus in on methods to identify and filter FFPE artifacts in sequencing data (DNA and RNA), the analysis of CLL and DLBCL cancer data, as well as the detection of somatic structural variants and gene fusions for the ICGC PCAWG project.

a. Lee, R. S., Stewart, C., Carter, S. L., Ambrogio, L., Cibulskis, K., Sougnez, C., et al. (2012). A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. The Journal of Clinical Investigation, 122(8), 2983–2988. PubMed Central PMCID: PMC3408754 ; PubMed PMID: 22797305
b. Crompton, B. D., Stewart, C., Taylor-Weiner, A., Alexe, G., Kurek, K. C., Calicchio, M. L., et al. (2014). The Genomic Landscape of Pediatric Ewing Sarcoma. Cancer Discovery. Nov;4(11):1326-41. PubMed PMID: 25186949
c. Cancer Genome Atlas Research Network. (2014). Integrated genomic characterization of papillary thyroid carcinoma. Cell, 159(3), 676–690. PubMed Central PMCID: PMC4243044
Cancer Genome Atlas Research Network. (2017) Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell. 32(2):185-203. PubMed PMID: 28810144.
d. Chapuy B., Stewart C., Dunford A.J., Kim J., et al. (2018), Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nature Medicine 24(5):679-690. PMID: 29713087

List of Published Work (excluding 31 physics publications not indexed by pubmed).
https://www.ncbi.nlm.nih.gov/sites/myncbi/1v1-Hv-gosqkc/bibliography/45000188/public/?sort=date&direction=ascending

D. Additional Information: Research Support and/or Scholastic Performance