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: Wendl, Michael

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

POSITION TITLE: Assistant Professor

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

INSTITUTION AND LOCATION	DEGREE (if applicable)	Completion Date MM/YYYY	FIELD OF STUDY
Washington University, St. Louis, MO	BS	05/1989	
Washington University, St. Louis, MO	MS	05/1990	
Washington University, St. Louis, MO	SCD	05/1994	

A. Personal Statement

This proposal outlines an ambitious project whose end-goal is the discovery of a large class of structural variants (translocations, complex SVs, etc) and establishment of their associations with various complex human diseases. Given the enormous data to which investigators will have access and the coupling with the advanced analytical/statistical methods, the work outlined in this proposal is very likely to ultimately yield important, clinically-relevant advancements in human disease. A substantial portion of this project will depend upon sophisticated mathematical, statistical, and computer science methodologies applied to analysis of genomic data, including SV detection, association/hypothesis testing, pathway/network analysis, and information systems. My primary role will be to contribute to and to help manage this aspect of the work. My abilities on a technical level to ensure its success in this regard are affirmed by mathematical and statistical contributions over a broad spectrum: in (1) bioinformatics, e.g. Phred, BreakDancer, MuSiC, PathScan, etc. (2) applied mathematics and statistics, e.g. coverage theory, and (3) pure mathematics, e.g. combinatorial probability, random variable collisions, and differential equations. Coupled with the expertise of the other investigators, we believe we can fully deliver on what we are proposing, and furthermore, that this work will indeed be an important foundation for the future of medicine and the diagnosis and treatment of complex human disease.

1. Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, Wendl MC, McMichael JF, Schmidt HK, Yellapantula V, Miller CA, Ozenberger BA, Welch JS, Link DC, Walter MJ, Mardis ER, Dipersio JF, Chen F, Wilson RK, Ley TJ, Ding L. Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nat Med. 2014 Dec;20(12):1472-8. PubMed PMID: 25326804; PubMed Central PMCID: PMC4313872.
2. Ding L, Raphael BJ, Chen F, Wendl MC. Advances for studying clonal evolution in cancer. Cancer Lett. 2013 Nov 1;340(2):212-9. PubMed PMID: 23353056; PubMed Central PMCID: PMC3783624.
3. Dees ND, Zhang Q, Kandoth C, Wendl MC, Schierding W, Koboldt DC, Mooney TB, Callaway MB, Dooling D, Mardis ER, Wilson RK, Ding L. MuSiC: identifying mutational significance in cancer genomes. Genome Res. 2012 Aug;22(8):1589-98. PubMed PMID: 22759861; PubMed Central PMCID: PMC3409272.
4. Wendl MC, Korf I, Chinwalla AT, Hillier LW. Automated processing of raw DNA sequence data. IEEE Eng Med Biol Mag. 2001 Jul-Aug;20(4):41-8. PubMed PMID: 11494768.

B. Positions and Honors

Positions and Employment

1994 - 2001	Research Associate, Washington University McDonnell Genome Institute, St. Louis, MO
2001 - 2005	Research Instructor, Washington University McDonnell Genome Institute, St. Louis, MO
2005 - 2014	Research Assistant Professor, Washington University McDonnell Genome Institute, St. Louis,

	MO
2010 -	Assistant Professor, Washington University Mathematics Dept, St. Louis, MO
2014 -	Assistant Professor, Washington University McDonnell Genome Institute, St. Louis, MO

Other Experience and Professional Memberships

Honors

1992	Graduate Student Fellow Award, American Institute of Aeronautics and Astronautics

C. Contribution to Science

1. I have contributed a number of bioinformatic and statistical tools/techniques that have become widely-used and well-established in the genomic sciences. My initial work was with Phil Green on PHRED, the analog-to-digital converter (Fast Fourier Transform and numerical optimization) that became the standard base-caller for Sanger sequencing. I have since contributed to processing tools (e.g. MuSiC), detectors (e.g. PolyScan, BreakDancer), and analysis tools (e.g. PathScan). Many of the other specialized statistical tests I have developed, for example, for allelic imbalance analysis, have appeared as part of a larger biological publication, rather than a standalone paper.

    a. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 1998 Mar;8(3):175-85. PubMed PMID: 9521921.

    b. Chen K, McLellan MD, Ding L, Wendl MC, Kasai Y, Wilson RK, Mardis ER. PolyScan: an automatic indel and SNP detection approach to the analysis of human resequencing data. Genome Res. 2007 May;17(5):659-66. PubMed PMID: 17416743; PubMed Central PMCID: PMC1855178.

    c. Wendl MC, Wallis JW, Lin L, Kandoth C, Mardis ER, Wilson RK, Ding L. PathScan: a tool for discerning mutational significance in groups of putative cancer genes. Bioinformatics. 2011 Jun 15;27(12):1595-602. PubMed PMID: 21498403; PubMed Central PMCID: PMC3106187.

    d. Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, Wendl MC, McMichael JF, Schmidt HK, Yellapantula V, Miller CA, Ozenberger BA, Welch JS, Link DC, Walter MJ, Mardis ER, Dipersio JF, Chen F, Wilson RK, Ley TJ, Ding L. Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nat Med. 2014 Dec;20(12):1472-8. PubMed PMID: 25326804; PubMed Central PMCID: PMC4313872.

2. Much of my applied mathematics work has focused on quantifying various molecular biology processes, e.g. mapping and sequencing DNA, requirements and power analysis for detecting various kinds of features, etc. This work revealed some non-intuitive surprises in designing post-HGP projects that subsequent experience has confirmed, for example that sequencing genomes of an individual for the purpose of detecting variation would need to be substantially higher than the 10X HGP standard (e.g. >30X, which is now standard), but that sequencing large populations to find rare variants is optimized for much lower depths, around 4x (a design subsequently used by the 1000 Genomes Project). Additional work has described optimization for many other scenarios, e.g. for detecting indels and for metagenomic sequencing of microbe communities.

    a. Wendl MC. A general coverage theory for shotgun DNA sequencing. J Comput Biol. 2006 Jul-Aug;13(6):1177-96. PubMed PMID: 16901236.

    b. Wendl MC, Wilson RK. Aspects of coverage in medical DNA sequencing. BMC Bioinformatics. 2008 May 16;9:239. PubMed PMID: 18485222; PubMed Central PMCID: PMC2430974.

    c. Wendl MC, Wilson RK. The theory of discovering rare variants via DNA sequencing. BMC Genomics. 2009 Oct 20;10:485. PubMed PMID: 19843339; PubMed Central PMCID: PMC2778663.

    d. Wendl MC, Kota K, Weinstock GM, Mitreva M. Coverage theories for metagenomic DNA sequencing based on a generalization of Stevens' theorem. J Math Biol. 2013 Nov;67(5):1141-61. PubMed PMID: 22965653; PubMed Central PMCID: PMC3795925.

3. My work in pure mathematics has been in 2 specific areas: partial differential equations (PDEs) and combinatorial probability. In the former, I have addressed some long-standing PDEs governed primarily by

the linear 2nd-order diffusion operator, including the Couette problem first discussed by GI Taylor in 1923. These analyses are closely applicable to fluid physics, heat transfer, etc. given the connection of the operator to physical diffusion processes. For the latter, I have focused on combinatorial problems that depend upon set-partitioning, including the collisions of random variables. This class of problems is widely applicable to peer-to-peer searching, physical collision and coverage processes, etc.

a. Wendl MC. General solution for the Couette flow profile. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics. 1999 Nov;60(5 Pt B):6192-4. PubMed PMID: 11970531.
b. Wendl MC. Mathematical analysis of coaxial disk cellular shear loading devices. The Review of scientific instruments. 2001; 72(11):4212-4217.
c. Wendl MC. Collision probability between sets of random variables. Statistics & probability letters. 2003; 64(3):249-254.

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

Ongoing Research Support

1R01CA178383-01A1, NCI
Ding, Li (PI)
09/08/14-08/31/17
Virus discovery and characterization in large-scale cancer sequencing data
Role: KP

CA172652-02, NCI
Chen, Ken (PI)
04/01/13-03/31/17
Delineating Heterogeneous Structural Complexity in Cancer Genomes
Role: KP

Completed Research Support

5R01HG005690-04, NHGRI
Raphael, Ben (PI)
01/01/11-12/31/15
Computational approaches for structural variation studies in genomes
Role: KP