## 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: William G. Fairbrother

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

POSITION TITLE: Professor

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 |
|---|---|---|---|
| Oberlin College (Oberlin, OH) | B.S. | 05/90 | Chemistry |
| Columbia University (New York City, NY) | M.A./M.Ph. | 10/96 | Biology |
| Columbia University (New York City, NY) | Ph.D. | 12/00 | Biology |
| Massachusetts Institute of Technology (Cambridge, MA) | Post-doctoral Fellow | 06/05 | Computational Biology |

### A. Personal Statement

I am a Professor in the MCB Department at the Warren Alpert School of Medicine at Brown University. I received my Ph.D. from the Department of Biological Sciences at Columbia University in 2000. I subsequently completed a postdoctoral fellowship in computational biology in lab of Christopher Burge and the lab of Nobel Laureate, Phillip Sharp, at Massachusetts Institute of Technology where I was a PhARMA Informatics Fellow. I developed the first computational screen for splicing enhancer elements (RESCUE-ESE, *Science* 2002). In 2005, I joined the faculty in the MCB Department at Brown University. In 2006, I became a member of the Center for Computational Molecular Biology (CCMB). I have continued my interest in building tools to predict which variants or disease mutations alter the processing of genes. We entered and won the CLARITY challenge – a personalized medicine contest that was held to develop analysis pipelines for personal genome data. I serve as Guest Editor of *PLoS Computational Biology* and have served on the program committees of ISMB- International Society of Computational Biology and also for the RNA society. I serve as an ad hoc member for several NIH study sections and am currently a permanent member of the GCAT section. I am currently one of seven Hassenfeld faculty scholars at Brown University overseeing genomic analysis for three state-wide child health initiatives (including autism initiative). My laboratory has ongoing collaborations with investigators throughout the Brown community to study transcriptome and spliceosome changes in a variety of conditions. We have established protocols for RNA-seq data analysis and assays to define splicing efficiency, accuracy, and spliceosome assembly. We developed the first massively parallel splicing assay (MaPSy) to screen variants (that occur in genomic sequence) for effects on splicing. The laboratory has expertise in integrating population genetic data, transcriptomic and functional genomic data into high throughput splicing analysis pipelines. We have expertise in analyzing large public and private variant datasets derived from genotyped patient cohorts such as *de novo* mutations in autism (Simon's Foundation), archaic variants (Neanderthal/Denisovan) and rare variants from the Geisinger patient cohort. In conclusion, we have developed numerous RNA genomic protocols and approaches involving multiplexed oligonucleotide pools, novel adaptions of NGS sequencing and RNA methodology. Many of these tools (e.g. MaPSY and using lariat reads to match branch points have been widely adopted in the field.

### B. Positions and Honors

**Positions and Employment**

1990 – 1991 Science Writer, Chemical Business, NY

1991 – 1994 Research Associate, Scientist's Institute for Public Information, NY
1994 – 2000 Ph.D. Candidate, Columbia University, NY

| | |
|---|---|
| 1994 – 1996 | Teaching Assistant/Computer Consultant |
| 2001 – 2003 | Post-doctoral Associate, Center for Cancer Research, MIT, Cambridge, MA |
| 2003 – 2005 | PhRMA Informatics Fellow, Center for Cancer Research, MIT, Cambridge, MA |
| 2005 – 2011 | Assistant Professor, Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, RI |
| 2006 – present | Member, Center for Computational Molecular Biology, Brown University, Providence, RI |
| 2006 – present | Member, Center for Genomics and Proteomics, Brown University, Providence, RI |
| 2011 – present | Member, Center for Biomedical Engineering, Brown University, Providence, RI |
| 2011 – present | Associate Professor, Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, RI |
| 2011 – present | Executive Committee, Center for Computational Molecular Biology, Brown University, Providence, RI |
| 2012 – 2013 | Visiting Fellow, Clare Hall, University of Cambridge, Cambridge, UK |
| 2013 – 2017 | Director of Graduate Studies, Center for Computational Molecular Biology, Brown University, Providence, RI |
| 2015 - present | Hassenfeld Faculty Scholar |
| 2017 – present | Full Professor, Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, RI |

**Other Experience and Professional Memberships**

| | |
|---|---|
| 2013 | Life Member, Clare Hall, University of Cambridge, Cambridge, UK |

**Honors**

| | |
|---|---|
| 2000 | McGregor Teaching Award |
| 2005 | Richard Salomon Award |
| 2005 | UTRA Summer Fellowship Award |
| 2006 | UTRA Summer Fellowship Award |
| 2007 | CCMB Scholarship Innovator Award |
| 2007 | UTRA Summer Fellowship Award |
| 2008 | UTRA Summer Fellowship Award |
| 2008 | CCMB Research Award “Incorporating Statistical Mechanics in Motif Finding” |
| 2009 | Research Seed Fund Award |
| 2012 | Winner – CLARITY Challenge |
| 2012 | Selected visiting fellow, Clare Hall, University of Cambridge |
| 2013 | Elected Life Member, Clare Hall, University of Cambridge |
| 2015 | Selected Hassenfeld Faculty Scholar |
| 2019 | Permanent member, GCAT NIH study section |

## C. Contributions to Science

### 1. First Computational Approach to Identifying Splicing Elements

The availability of human genomic sequence drove a series of computational approaches to identifying cis-regulatory elements that drive gene expression. During my postdoc in the lab of Chris Burge and Phil Sharp, we developed the first computational method for finding exonic splicing enhancers (ESE) based on their enrichment around sub-optimal splice sites. This resulted work the creation of online annotation tools, studies of cis-element evolution across species and further efforts to discover other types of splicing elements such as intronic enhancers and silencers. Along these lines we have developed model free methods of identifying functional elements and mutation prediction utilizing positional distribution.

a) **Fairbrother, W.G.**, Yeh, R.F., Sharp, P.A., and Burge, C.B. (2002). Predictive identification of exonic splicing enhancers in human genes. Science *297(5583)*, 1007-1013. (PMID: 12114529)

b) **Fairbrother, W.G.**, Holste, D., Burge, C.B., and Sharp, P.A. (2004a). Single nucleotide polymorphism-based validation of exonic splicing enhancers. PLoS Biol. *2(9)*, e268. (PMID: 15340491; PMCID: PMC514884)

c) **Fairbrother, W.G.**, Yeo, G.W., Yeh, R., Goldstein, P., Mawson, M., Sharp, P.A., and Burge, C.B. (2004b). RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. Nucleic Acids Res. *32*, W187-190. (PMID: 15215377; PMCID: PMC441531)

d) Lin CL, Taggart AJ, Lim KH, Cygan KJ, Ferraris L, Creton R, Huang YT, **Fairbrother WG**. RNA structure replaces the need for U2AF2 in splicing. Genome Res. 2016 Jan;26(1):12-23. doi: 10.1101/gr.181008.114. Epub 2015 Nov 13. (PubMed PMID: 26566657; PubMed Central PMCID: PMC4691745)

2. **First Large-Scale Discovery and Mapping of Human Branch points** Human pre-mRNAs are extensively processed. A typical transcript contains about 10 introns that are removed by splicing. Each splicing event produces an exon junction and an excised intron lariat. The intron lariat is an unusual branched RNA. Not only is the location of the branch highly predictive of alternative splicing but the lariat itself can function as a non-coding RNA (e.g. an expression vector for snoRNAs and microRNAs and a potential modifier of chromatin). We developed techniques to profile lariats and map branchpoints from deep sequencing reads. Our technique has been incorporated into NGS processing software (e.g. Lasso) and also been adopted by others (Pleis lab, Mattick lab). We are continuing this work as an active line of research in the lab.
   a) Taggart, A.J., DeSimone, A.M., Shih, J.S., Filloux, M.E., and **Fairbrother, W.G.** (2012). Large-scale mapping of branchpoints in human pre-mRNA transcripts in vivo. Nat. Struct. Mol. Biol. *19(7)*, 719-721. (PMID: 22705790; PMCID: PMC3465671)
   b) Taggart AJ, Lin CL, Shrestha B, Heintzelman C, Kim S, **Fairbrother WG**. Large-scale analysis of branchpoint usage across species and cell lines. Genome Res. 2017 Apr;27(4):639-649. doi: 10.1101/gr.202820.115. Epub 2017 Jan 24. (PubMed PMID: 28119336; PubMed Central PMCID: PMC5378181)

3. **Developing Protocols for Personalized Medicine** my lab has developed computational tools to predict the effects of genetic variants on the processing of RNA. In this regard, we have practical experience with clinical applications, specifically, the problem of determining which variants identified in clinical sequencing data might be the causal alleles. In 2012, we were a member of the team that won a personalized medicine contest involving whole-genome sequencing and exome sequencing data – the "CLARITY challenge." We have utilized these technologies to identify variants that cause human disease.
   a) Brownstein CA, Beggs AH, Homer N, Merriman B, Yu TW, Flannery KC, DeChene ET, Towne MC, Savage SK, Price EN, Holm IA, Luquette LJ, Lyon E, Majzoub J, Neupert P, McCallie D Jr, Szolovits P, Willard HF, Mendelsohn NJ, Temme R, Finkel RS, Yum SW, Medne L, Sunyaev SR, Adzhubey I, Cassa CA, de Bakker PI, Duzkale H, Dworzyński P, **Fairbrother W**, Francioli L, Funke BH, Giovanni MA, Handsaker RE, Lage K, Lebo MS, Lek M, Leshchiner I, MacArthur DG, McLaughlin HM, Murray MF, Pers TH, Polak PP, Raychaudhuri S, Rehm HL, Soemedi R, Stitziel NO, Vestecka S, Supper J, Gugenmus C, Klocke B, Hahn A, Schubach M, Menzel M, Biskup S, Freisinger P, Deng M, Braun M, Perner S, Smith RJ, Andorf JL, Huang J, Ryckman K, Sheffield VC, Stone EM, Bair T, Black-Ziegelbein EA, Braun TA, Darbro B, DeLuca AP, Kolbe DL, Scheetz TE, Shearer AE, Sompallae R, Wang K, Bassuk AG, Edens E, Mathews K, Moore SA, Shchelochkov OA, Trapane P, Bossler A, Campbell CA, Heusel JW, Kwitek A, Maga T, Panzer K, Wassink T, Van Daele D, Azaiez H, Booth K, Meyer N, Segal MM, Williams MS, Tromp G, White P, Corsmeier D, Fitzgerald-Butt S, Herman G, Lamb-Thrush D, McBride KL, Newsom D, Pierson CR, Rakowsky AT, Maver A, Lovrečić L, Palandačić A, Peterlin B, Torkamani A, Wedell A, Huss M, Alexeyenko A, Lindvall JM, Magnusson M, Nilsson D, Stranneheim H, Taylan F, Gilissen C, Hoischen A, van Bon B, Yntema H, Nelen M, Zhang W, Sager J, Zhang L, Blair K, Kural D, Cariaso M, Lennon GG, Javed A, Agrawal S, Ng PC, Sandhu KS, Krishna S, Veeramachaneni V, Isakov O, Halperin E, Friedman E, Shomron N, Glusman G, Roach JC, Caballero J, Cox HC, Mauldin D, Ament SA, Rowen L, Richards DR, San Lucas FA, Gonzalez-Garay ML, Caskey CT, Bai Y, Huang Y, Fang F, Zhang Y, Wang Z, Barrera J, Garcia-Lobo JM, González-Lamuño D, Llorca J, Rodriguez MC, Varela I, Reese MG, De La Vega FM, Kiruluta E, Cargill M, Hart RK, Sorenson JM, Lyon GJ, Stevenson DA, Bray BE, Moore BM, Eilbeck K, Yandell M, Zhao H, Hou L, Chen X, Yan X, Chen M, Li C, Yang C, Gunel M, Li P, Kong Y, Alexander AC, Albertyn ZI, Boycott KM, Bulman DE, Gordon PM, Innes AM, Knoppers BM, Majewski J, Marshall CR, Parboosingh JS, Sawyer SL, Samuels ME, Schwartzentruber J, Kohane IS, Margulies DM. An international effort towards developing standards for best practices in analysis, interpretation and reporting of clinical genome sequencing results in the CLARITY Challenge. Genome Biol. 2014 Mar 25;15(3):R53. doi: 10.1186/gb-2014-15-3-r53. (PubMed PMID: 24667040; PubMed Central PMCID: PMC4073084.)

b) Ceyhan-Birsoy O, Agrawal PB, Hidalgo C, Schmitz-Abe K, DeChene ET, Swanson LC, Soemedi R, Vasli N, Iannaccone ST, Shieh PB, Shur N, Dennison JM, Lawlor MW, Laporte J, Markianos K, **Fairbrother WG**, Granzier H, Beggs AH. Recessive truncating titin gene, TTN, mutations presenting as centronuclear myopathy. Neurology. 2013 Oct 1;81(14):1205-14. doi: 10.1212/WNL.0b013e3182a6ca62. Epub 2013 Aug 23. (PubMed PMID: 23975875; PubMed Central PMCID: PMC3795603.)

c) Pfarr, N., Prawitt, D., Kirschfink, M., Schroff, C., Knuf, M., Habermehl, P., Mannhardt, W., Zepp, F., **Fairbrother, W.G.**, Loos, M., Burge, C.B., and Pohlenz, J. (2005). Linking C5 deficiency to an exonic splicing enhancer mutation. J. Immunol. *174(7)*, 4172-4177. (PMID: 15778377; PMCID: Not available)

d) Waugh JL, Celver J, Sharma M, Dufresne RL, Terzi D, Risch SC, **Fairbrother WG**, Neve RL, Kane JP, Malloy MJ, Pullinger CR, Gu HF, Tsatsanis C, Hamilton SP, Gold SJ, Zachariou V, Kovoor A. Association between regulator of G protein signaling 9-2 and body weight. PLoS One. (2011);6(11):e27984. doi: 10.1371/journal.pone.0027984. Epub 2011 Nov 23. (PubMed PMID: 22132185; PubMed Central PMCID: PMC3223194.)

4. **Developing Platforms for High Throughput Biochemical Assays**. We have developed highly multiplex nucleic acid binding and splicing assays. Utilizing deep sequencing technology and synthetic library preparations we have developed assay and analysis tools to measure thousands of binding readouts. This high throughput implementation of traditional binding and splicing assays combine the flexibility of biochemistry and the survey breadth of genomic analysis.

a) Soemedi R, Cygan KJ, Rhine CL, Wang J, Bulacan C, Yang J, Bayrak-Toydemir P, McDonald J, Fairbrother WG. Pathogenic variants that alter protein code often disrupt splicing. Nat Genet. 2017 Jun;49(6):848-855. doi: 10.1038/ng.3837. Epub 2017 Apr 17. (PubMed PMID: 28416821)

b) Ferraris, L., Stewart, A.P., Gemberling, M.P., Reid, D.C., Lapadula, M.J., Thompson, W.A., and Fairbrother, W.G. (2011). High-throughput mapping of protein occupancy identifies functional elements without the restriction of a candidate factor approach. Nucleic Acids Res. 39(6), e33. (PMID: 21169336; PMCID: PMC3064794)

c) Reid, D.C., Chang, B.L., Gunderson, S.I., Alpert, L., Thompson, W.A., and **Fairbrother, W.G.** (2009). Next-generation SELEX identifies sequence and structural determinants of splicing factor binding in human pre-mRNA sequence. RNA 15(12), 2385-2397. (PMID: 19861426; PMCID: PMC2779669)

d) Tantin, D., Gemberling, M., Callister, C., and **Fairbrother, W.G.** (2008). High-throughput biochemical analysis of in vivo location data reveals novel distinct classes of POU5F1(Oct4)/DNA complexes. Genome Res. 18(4), 631-639. (PMID: 18212089; PMCID: PMC2279250)

5. **Computational and Experimental Approaches to Discover Variants that affect pre-mRNA splicing** We have developed a variety of online tools and the high throughput assay (described above) that are designed to identify causal variants in sequencing data. Specifically, these tools detect variants that disrupt splicing.

a) Watkins, K.H., Stewart, A., and **Fairbrother, W.** (2009). A rapid high-throughput method for mapping ribonucleoproteins (RNPs) on human pre-mRNA. J. Vis. Exp. 34, pii1622. (PMID: 19956082; PMCID: PMC3152247)

b) Lim, K.H., and **Fairbrother, W.G.** (2012). Spliceman--a computational web server that predicts sequence variations in pre-mRNA splicing. Bioinformatics 28, 1031-1032. (PMID: 22328782; PMCID: PMC3315715)

c) Soemedi R, Vega H, Belmont JM, Ramachandran S, **Fairbrother WG**. Genetic variation and RNA binding proteins: tools and techniques to detect functional polymorphisms. Adv Exp Med Biol. 2014;825:227-66. doi: 10.1007/978-1-4939-1221-6_7. (PubMed PMID: 25201108.)

d) Soemedi R, Cygan KJ, Rhine CL, Wang J, Bulacan C, Yang J, Bayrak-Toydemir P, McDonald J, **Fairbrother WG**. Pathogenic variants that alter protein code often disrupt splicing. Nat Genet. 2017 Jun;49(6):848-855. doi: 10.1038/ng.3837. Epub 2017 Apr 17. (PubMed PMID: 28416821.)

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

**Ongoing Research Support**

3R01GM105681-08S1 FAIRBROTHER, WILLIAM G(PI) 7/16/14 – 1/31/23
"Supplement to: A genomic approach to studying the life cycle of intron lariats"

R01 GM105681 Fairbrother (PI) 7/16/14 – 1/31/23
“A Genomic Approach to Studying the Life Cycle of Intron Lariats”

R56 MH127844 FAIRBROTHER, WILLIAM G (co-PI) 2021 -2022
KNOWLES, DAVID ARTHUR co-PI)
Principal Investigator(s)/ Project Leader(s) $772,112
“Fine-mapping psychiatric disease variants that affect post-transcriptional gene regulation”

The major goals of this project are to: (1) map branchpoints in human transcripts; (2) map branchpoints around known targets of heterogeneous nuclear ribonucleoprotein (hnRNP) protein-binding and regulatory sites; and (3) discover determinants of branchpoint turnover and regulation.
Role: PI

R01 GM127472 Fairbrother (PI) 8/23/18 – 4/30/22
"Discovering Splicing Defects in Human Genes"

The major goals of this project are to screen human variants for their ability to disrupt splicing and to make further improvements to the high throughput assay. All known variants within the vicinity of splice sites in a set of actionable genes' will be analyzed. The GHS patient panel will be used for validating predictions in human tissues.

**Completed Research Support**

Simons Foundation Fairbrother (PI) 9/1/15-9/1/17
SFARI Award
"Use of High-throughput Splicing Assays to Prioritize Autism Gene Candidates"
The goal of this project is to better map splicing elements in SSC genes and prioritize candidate causal variants in SSC sequencing data by analyzing processing phenotype of variants.
Role: PI