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: Yi Yin

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

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. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE (if applicable)	END DATE MM/YYYY	FIELD OF STUDY
Beijing Normal University, Beijing, China	B.S.	07/2009	Biotechnology
Duke University, Durham, NC	M.S.	09/2015	Statistical Science
Duke University, Durham, NC	Ph.D.	09/2015	Genetics and Genomics
Harvard University, Cambridge, MA	Postdoctoral	07/2016	Postdoctoral Fellowship in Chemistry and Chemical Biology
University of Washington, Seattle, WA	Postdoctoral	02/2020	Postdoctoral Fellowship in Genome Sciences

A. Personal Statement

Homologous recombination (HR) is essential in embryonic development, meiosis and tumorigenesis. My long-term goal is to understand HR and resulting genome instability in a fully probabilistic manner. Given one's genotype, can we predict: 1) which genome regions are fragile; 2) what (epi)genomic contexts regulate DNA breakage; 3) how mutations and expression levels of DNA repair genes affect repair processes; 4) what consequences HR and resulting rearrangements have from a cell to an individual? To achieve this goal, I had interdisciplinary training in 1) studying fundamental HR mechanism, 2) Bayesian statistical modeling, and 3) technology development.

For my doctoral work in Tom Petes lab, I generated the first genome-wide maps of mitotic HR events in any organism by analyzing loss-of-heterozygosity tracts in reciprocal daughter cell clones in the wild-type and mutants[1-2]. By global mapping of heteroduplex DNA, an HR intermediate, I dissected all the eight strands involved in HR and provided new insights into the complexity of its mechanism[3]. I also completed a concurrent MS degree in Statistical Science, where I took 10 PhD-level courses, passed First-Year Exam for Duke Statistical Science PhD students, and completed an MS thesis in predictive modeling of HR hotspots. I am thus mathematically mature to tackle statistical questions derived from large mapping datasets.

For my postdoc, I trained with Sunney Xie and Jay Shendure on improving chemistry and throughput of single-cell assays. Combining strengths from both labs, I developed sci-L3, which enables high-coverage single-cell whole-genome sequencing with throughput of 1M cells and generalizability to other assays such as target-seq and DNA/RNA co-assay[4]. I applied sci-L3 to map meiotic crossover and chromosome segregation in >10,000 germ cells and discovered for the first time whole-genome equational (mitotic-like) segregation during meiosis I in both fertile and infertile mice.

Recently, I have further expanded sci-L3 toolset which set the technical cornerstones for our current research. We also contributed to COVID-19 testing and sequencing at UCLA, where I developed V-seq, an inexpensive, fast, and scalable method that performs targeted SARS-CoV-2 genome sequencing. The approach can be generalized to other applications of targeted RNA-seq.

The research in our lab can be a bridge between new technologies and new insights into HR mechanisms, which builds solidly upon the foundations of my past work, as it leverages a combination of developing advanced single-cell sequencing technologies and complementary computational tools to tackle pressing questions about causes and consequences of genome instability central to the understanding of cancer initiation and treatment.

1. **Yin Y** & Petes TD. *PLoS Genet.* 9, e1003894 (2013).
2. **Yin Y** & Petes TD. *PLoS Genet.* 11, e1005026 (2015).
3. **Yin Y**, Dominska M, Yim E & Petes TD. *Elife* 6, (2017).
4. **Yin Y**†, Jiang Y, Lam KG, Berletch JB, Disteche CM, Noble WS, Steemers FJ, Camerini-Otero RD, Adey AC, Shendure JA†. *Mol. Cell*. 76(4):676-690.e10 (2019) †Co-corresponding authors.

B. Positions and Honors

Positions and Employment

2006 - 2008	Undergraduate Student, Yu-Sheng Cong Lab, Beijing Normal University, Beijing, China
2008	Summer Student, Xiaochen Wang Lab, National Institute of Biological Sciences, Beijing, China
2009	Undergraduate Student, Stephen Cohen Lab, TLL Temasek Life Sciences Laboratory, National University of Singapore, Singapore
2010 - 2015	Doctoral Student, Tom Petes lab, Duke University - University Program in Genetics and Genomics
2013 - 2015	Master Student, Sayan Mukherjee Group, Duke University - Department of Statistical Science
2015 - 2016	Damon Runyon Postdoctoral Fellow, Sunney Xie lab, Harvard University - Department of Chemistry and Chemical biology
2016 - 2020	Damon Runyon Postdoctoral Fellow (Part-Time Lecturer), Jay Shendure Lab, University of Washington - Department of Genome Sciences
2020 -	Assistant Professor, University of California, Los Angeles - Department of Human Genetics

Other Experience and Professional Memberships

2011 - 2013	Certificate in Computational Biology and Bioinformatics, Duke University
2012	Teaching Assistant, Genetic and Genomics Solutions to Biological Problems (*C.elegans* session, led journal club discussions), Duke University
2013	Teaching Assistant, Genetic and Genomics Solutions to Biological Problems (Computational Approaches to Identify Protein Binding Sites session, tutored Perl programming), Duke University
2013 - 2015	Member, Genetics Society of America
2013 -	Students supervised: Sarah Jaslow (Duke MGM), Eunice Yim (Duke Biology), Wenting Cai (Harvard CCB), Shawn Fayer (UW GS)
2015 -	Ad hoc reviewer for Molecular Cell, Nature Genetics, PNAS, Nucleic Acids Research, Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis, FEMS Yeast Research, Mutagenesis, Applied Microbiology and Biotechnology
2017	Lecturer (GENOME541, Bayesian Statistics session), University of Washington
2019	Organizer of Career Development Workshops, FASEB, Genetic Recombination and Genome Rearrangements, Steamboat Springs, CO
2020 -	Member, American Association for Cancer Research
2020 -	Member, Molecular Biology Institute, UCLA

Honors, Awards, Invited talks

2006, 2008	First-class scholarship (top 3%, merit-based), Beijing Normal University, China.
2007	Second-class scholarship (top 6%, merit-based), Beijing Normal University, China.
2013	Chairman's Travel Awards, Duke University Medical Center.
2015	Certificate of Outstanding Contribution in Reviewing, Mutation Research.
2016 - 2019	Damon Runyon Postdoctoral Fellowship Award, Damon Runyon Cancer Research Foundation
2017	*Invited talk,* Tom Petes 70th Birthday Symposium, Duke University (Durham, NC).
2018	Outstanding Poster Presentation Award, Meiosis - Gordon Research Seminars (GRS).
2018	*Invited talk*, Illumina Sequencing Symposium (Seattle, WA).
2019	*Invited talk*, Human Genetics Faculty Search Symposium, UCLA (Los Angeles, CA).
2019	*Invited talk*, Molecular Biology Program, Sloan Kettering Institute, MSKCC (New York, NY).
2019	*Invited talk*, Stadtman Tenure-Track Investigator Candidate Seminar, NIA (Baltimore, MD).
2019	*Invited talk*, Stadtman Tenure-Track Investigator Candidate Seminar, NIEHS (Durham, NC).
2019	Finalist (22/281 applicants), Career Award at the Scientific Interface, Burroughs Wellcome Fund.
2020	*Invited talk*, MBI Faculty Seminar, UCLA (Los Angeles, CA).
2020	*Invited talk*, SoCal Genome Stability Symposium (San Diego, CA); postponed to 2021.
2020	*Invited talk,* DNA Dynamics Webinar Series, Columbia University (New York, NY).
2020	*Invited talk,* COVID-19 Basic, Translational and Clinical Research Seminar Series, UCLA (Los Angeles, CA).

2020 *Invited talk,* Genomic Stability and DNA repair, Virtual Keystone Symposia, USA
2020 *Invited talk,* Bioinformatics-Human Genetics Seminar, UCLA (Los Angeles, CA).
2020 - 2022 Damon Runyon - Dale F. Frey Award for Breakthrough Scientists.

C. Contribution to Science

1. **Genome-wide mapping of UV-induced mitotic HR events in yeast.** I performed the first global mapping of UV-induced mitotic HR events. UV radiation is known to generate single strand nicks. However, when I mapped UV-induced HR events in reciprocal daughter cell clones in hybrid diploid yeast, there is strong evidence that UV directly causes double strand breaks (DSB) in unreplicated chromosomes[a,b]. I followed up on the genetic requirements for UV-induced DSBs and provided mechanistic insights. We also found to our surprise that UV induces 7-fold more HR events than X-ray at an equitoxic dose. I then studied how UV lesions are processed to be highly recombinogenic and characterized the roles of nucleotide excision repair (NER), structural nucleases and post-replication repair. I found that unexcised UV dimers in NER-deficient cells strongly stimulate HR and that Mus81p is responsible for cleaving the stalled replication fork to generate the DSBs. I also found that Mms2p-mediated post-replication repair primarily promotes HR between sister chromatids and acts to suppress HR between homologs[c].

    a. St Charles J, Hazkani-Covo E, **Yin Y**, Andersen SL, Dietrich FS, Greenwell PW, Malc E, Mieczkowski P, Petes TD. High-resolution genome-wide analysis of irradiated (UV and γ-rays) diploid yeast cells reveals a high frequency of genomic loss of heterozygosity (LOH) events. *Genetics.* 2012; 190(4): 1267-84. PMID: 22267500 PMCID: PMC3316642.
    b. **Yin Y**, Petes TD. Genome-wide high-resolution mapping of UV-induced mitotic recombination events in *S. cerevisiae*. *PLoS Genet.* 2013; 9(10): e1003894. PMID: 24204306 PMCID: PMC3814309.
    c. **Yin Y**, Petes TD. Recombination between homologous chromosomes induced by unrepaired UV-generated DNA damage requires Mus81p and is suppressed by Mms2p. *PLoS Genet.* 2015; 11(3): e1005026. PMID: 25738287 PMCID: PMC4349867.

2. **Global analyses of spontaneous and UV-induced mitotic HR, genome rearrangements and aneuploidy in DNA repair mutants.** I examined spontaneous and UV-induced mitotic HR events in strains defective of various steps, including: 1) initial end resection pathways[a] (with Lorraine Symington lab), 2) extensive end resection pathways[b], 3) helicases (dissertation) and 4) a non-essential subunit of DNA polymerase (dissertation). I also built predictive models of HR hotspots using Bayesian stochastic search variable selection (MS thesis). The global analyses of HR provided insights not available in studies of induced DSB at a defined locus, including distinct types of events in different genetic mutants, and the non-random distributions of events associated with various DNA damage.

    a. Deng SK, **Yin Y**, Petes TD, Symington LS. Mre11-Sae2 and RPA Collaborate to Prevent Palindromic Gene Amplification. *Mol. Cell.* 2015; 60(3): 500-8. PMID: 26545079 PMCID: PMC4636734.
    b. **Yin Y**, Petes TD. The role of Exo1p exonuclease in DNA end resection to generate gene conversion tracts in *S. cerevisiae*. *Genetics.* 2014; 197(4): 1097-109. PMID: 24835424 PMCID: PMC4125386.

3. **Mapping of heteroduplex DNA, an HR intermediate, in mitotic HR events.** I generated the first genome-wide map of heteroduplex DNA formed in spontaneous and UV-induced mitotic HR events, which is an intermediate diagnostic of all the eight strands involved in HR. Based on simulation of all possible HR profiles predicted by “classic” HR models, I showed that these “classic” models of DSB repair are inadequate to explain several aspects of HR[a]. Qualitatively, we provided evidence that additional HR mechanisms, such as template switching during repair synthesis, branch migration of Holliday junctions, and Mlh1-independent mismatch correction, have to be evoked to account for complex HR profiles we observed.

    a. **Yin Y**, Dominska M, Yim E, Petes T. High-resolution mapping of heteroduplex DNA formed during UV-induced and spontaneous mitotic recombination events in yeast. *eLife.* 2017; 6. pii: e28069. PMID: 28714850 PMCID: PMC5531827.

4. **Developing new single-cell DNA sequencing and multiplexed CRISPR genome editing technologies.** Conventional methods for single cell sequencing are limited with respect to uniformity and throughput. I

developed 'sci-L3', a single cell sequencing method that combines combinatorial indexing ('sci-') and linear ('L') amplification. The sci-L3 method adopts a 3-level ('3') indexing scheme that minimizes amplification biases while enabling exponential gains in throughput. We demonstrate the generalizability of sci-L3 with single-cell whole genome sequencing (WGS), targeted sequencing, and a co-assay of the genome and transcriptome. We applied sci-L3-WGS to profile the genomes of >10,000 sperm and sperm precursors from F1 hybrid mice, mapping 86,786 crossovers and characterizing rare chromosome mis-segregation events in male meiosis, including instances of whole-genome equational (mitotic-like) chromosome segregation in both infertile, interspecific and fertile, intraspecific mice. This application has novelty in three aspects: 1) it provides a genetic tool for studying meiosis in infertile, inter-specific strains as traditional analysis of F2 mice requires fertility; 2) the throughput enables recovering rare cell population that has completed MeiosisI (MI) but not MII, a unique window where crossover and chromosome mis-segregation can be studied together; 3) we observed cell-autonomous equational segregation in MI for the first time in mammals. We anticipate that sci-L3 can be applied to fully characterize HR landscapes, to couple CRISPR perturbations and measurements of genome stability, and to other goals requiring high-throughput single cell sequencing[a]. As side projects, I also assisted in developing large-scale CRISPR genome editing in yeast[b] (with George Church lab) as well as profiling and predictive modeling of CRISPR repair outcomes in mammalian cell lines[c].

a. **Yin Y**[†], Jiang Y, Lam KG, Berletch JB, Disteche CM, Noble WS, Steemers FJ, Camerini-Otero RD, Adey AC, Shendure JA[†]. High-throughput single cell sequencing with linear amplification[#]. *Mol. Cell*. 2019; 76(4):676-690.e10. [†]Co-corresponding authors. PMID: 31495564 PMCID: PMC6874760.

b. Guo X, Chavez A, Tung A, Chan Y, Kaas C, **Yin Y**, Cecchni R, Lopez-Garnier S, Kelsic E, Schubert M, DiCarlo J, Collins JJ, Church GM. High-throughput creation and functional profiling of DNA sequence variant libraries using CRISPR–Cas9 in yeast. *Nat Biotechnol.* 2018; 36(6): 540-546. PMID: 29786095 PMCID: PMC5990468.

c. Chen W, McKenna A, Schreiber J, Haeussler M, **Yin Y**, Agarwal V, Noble WS, Shendure JA. Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. *Nucleic Acids Res.* 2019; pii: gkz487. 31165867 PMCID: PMC6735782.

5. **Developing new COVID-19 testing and sequencing technologies.** Conventional methods for viral genome sequencing largely use metagenomic approaches, or enrich for viral genomes by hybridization-based capture or amplicon sequencing with virus-specific PCR. These methods are costly, require extensive sample handling, and have limited throughput. I developed V-seq, an inexpensive, fast, and scalable method that performs targeted viral genome sequencing by multiplexing virus-specific reverse transcription (RT) primers. I designed densely tiled RT primers across the viral genome, with a subset of hexamers at the 3' end to minimize mis-priming from the abundant human rRNA repeats and RNA PolII transcriptome. We found that overlapping RT primers do not interfere, but rather act in concert to improve viral genome coverage in samples with low viral load. We anticipate that V-seq can be applied to investigate viral genome evolution in a cost-effective manner. More broadly, V-seq can be used in other applications of targeted RNA-seq[a]. I also co-supervised the epidemiology work for SARS-CoV-2 strains circulating in LA using V-seq[b], and assisted in developing Swab-seq as a high-throughput platform for massively scaled up SARS-CoV-2 testing[c].

a. Guo L, Boocock J, Tome JT, Chandrasekaran S, Hilt EE, Zhang Y, Sathe L, Li X, Luo C, Kosuri S, Shendure JA, Arboleda VA, Flint J, Eskin E, Garner OB, Yang S, Bloom JS[†], Kruglyak L[†], **Yin Y**[†]. Rapid cost-effective viral genome sequencing by V-seq. [†]Co-corresponding authors. *bioRxiv*, DOI: 10.1101/2020.08.15.252510.

b. Guo L[*], Boocock J[*], Hilt EE, Chandrasekaran S, Zhang Y, Munugala C, Sathe L, Alexander N, Arboleda VA, Flint J, Eskin E, Luo C, Yang S, Garner O, **Yin Y**[†], Bloom JS[†], Kruglyak L[†]. The genetic epidemiology of the Los Angeles COVID-19 outbreak. [†]Co-corresponding authors. *medRxiv*, DOI: 10.1101/2020.09.15.194712.

c. Bloom JS[†], Jones EM, Gasperini M, Lubock NB, Sathe L, Munugal C. Booeshaghi AS, Brandenberg OF, Guo L, Simpkins SW, Lin I, LaPierre N, Hong D, Zhang Y, Oland G, Choe BJ, Chandrasekaran S, Hilt EE, Butte M, Damoiseaux R, Cooper AR, **Yin Y**, Pachter L, Garner OB, Flint J, Eskin E, Luo C, Kosuri S[†], Kruglyak L[†], Arboleda VA[†]. Swab-Seq: A high-throughput platform for massively scaled up SARS-CoV-2 testing. [†]Co-corresponding authors. *medRxiv,* DOI: 10.1101/2020.08.04.20167874.

Complete List of Published Work in MyBibliography:
https://www.ncbi.nlm.nih.gov/myncbi/yi.yin.2/bibliography/public/

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

Ongoing Research Support

Damon Runyon - Dale F. Frey Award for Breakthrough Scientists, Damon Runyon Cancer Research Foundation 04/01/20-03/31/22

Global analysis of DNA break repair by single-cell sequencing

The goal of the project is to study genomic and evolutionary variation in DNA repair mechanisms with single cell sequencing technologies.

Role: PI

Completed Research Support

Damon Runyon Fellowship, Damon Runyon Cancer Research Foundation 01/01/16-12/31/19

Global analysis of DNA break repair by single-cell sequencing

The goal of this project is to develop state-of-the-art single cell DNA sequencing technology to examine how DNA repair mechanisms go awry and contribute to cancer initiation and progression.

Role: PI