Section 1:

CBC CHICAGO BIOMEDICAL CONSORTIUM

CBC Catalyst Award Proposal Title Page

Title of the proposal *(no more than 100 characters including white space)*

FUNCTIONS AND ORIGINS OF NEW COMPONENTS IN THE CCR4-NOT DEADENYLASE COMPLEX IN DROSOPHILA

Name, degree, title, institution, and contact information including the email address of each PI

PI #1 Manyuan Long, Ph.D., Edna K. Papazian Distinguished Service Professor, Department of Ecology and Evolution and the College, the University of Chicago, 1101 E 57th Street, Chicago, IL 60637. Email: mlong@uchicago.edu

PI #2 Jie Liang, Ph.D., Richard and Loan Hill Professor, Richard and Loan Hill Department of Biomedical Engineering (MC 063), the University of Illinois Chicago, 851 South Morgan Street, Chicago, IL 60607-7052. jliang@uic.edu

PI #3 Liming Li, Ph,D., Associate Professor, Department of Biochemistry and Molecular Genetics, Northwestern University, the Feinberg School of Medicine, 303 E. Superior Street, Chicago, IL 60611. limingli@northwestern.edu

Do you have any current or pending grant applications that potentially overlap with this application?
Check ONE:

YES ☐ **NO** ☒

If YES, please identify them.

(ENTER THE OVERLAP HERE)

Explain the overlap in the Biosketch section.

Does the proposed research involve animal subjects?	**Check ONE:**	**YES** ☐	**NO** ☒
Does the proposed research involve human subjects?	**Check ONE:**	**YES** ☐	**NO** ☒
Does the proposed research involve embryonic stem cells?	**Check ONE:**	**YES** ☐	**NO** ☒

(Note: if the data entered exceeds one page, it is acceptable to submit it as two pages.)

Section 2:

Summary. The Long lab has recently discovered that the X-linked gene, *Caf40*, which encodes a subunit of the essential CCR4-NOT deadenylase complex in eukaryotes, was recently duplicated twice through retroposition and the duplicated paralogs landed on two different autosomes with different expression patterns. Contrary to the conventional notion that similar gene duplicates are functionally redundant, these duplicated genes, *Poseidon* and *Zeus*, have acquired new functions in male fertility and development. It is unclear, however, how these newly duplicated *Caf40* paralogs interact with the parental CCR4-NOT complex and carry out their new, divergent roles in fertility and development. To investigate this, we propose to create CRISPR-Cas9 deletions of both genes to investigate their effects on development and male fertility. We also propose to use the AlphaFold system to predict protein-protein interactions of *Poseidon* and *Zeus* with CCR4-NOT subunits, and to construct their interaction networks with other partners. We will experimentally confirm the physical interactions of these predicted binders. We will also identify their protein interactomes using the co-immunoprecipitation coupled with mass spectrometry analysis. Results from this proposed study will shed light on how a conserved genetic system can evolve rapidly to give rise to new genes with distinct functions.

Our recent observations (Fig 1) have revealed a surprising process that young gene duplicates in *Drosophila* quickly have established essential or critically important functions in reproductive and developmental processes, challenging a general convention that important functions are encoded in conserved genes. To further understand the mechanism underlying this novel phenomenon, we propose an in-depth analysis of two such new duplicated genes derived from *Caf40*, *Poseidon* and *Zeus*, by investigating their effects on reproductive functions and testing their interactions and relationship with their parental CCR4-NOT complex. Our proposed study is understandably at a high risk for general funding sources due to the novelty of system. However, if succeeded, our study will bring out significant results shifting current paradigm of origin of novel biological functions. We are confident that we will be able to complete the two specific aims in this proposal.

Description of proposed research. It is generally accepted that essential functions are associated with conserved genes that are shared by many organisms, e.g. [9]. However, using RNAi-knockdown and CRISPR-knockout, we identified 138 recently evolved genes in *Drosophila melanogaster* that have acquired essential functions for reproduction and/or development [21]. These genes are young-aged from species-specific (200 thousand years ago, Kya) to subgenus-specific (<35 million years ago, Mya), with a high frequency of occurrence as shown in Figure 1A [10,24]. We have explored the underlying mechanisms at various levels from chromosomes, tissues, individuals to populations in order to understand how these species-specific or lineage-specific young genes have developed essential functions [8,11,18]. Given this potential paradigm-shifting opportunity, we propose to further tackle the important but little-known problem of how redundant duplicates become essential.

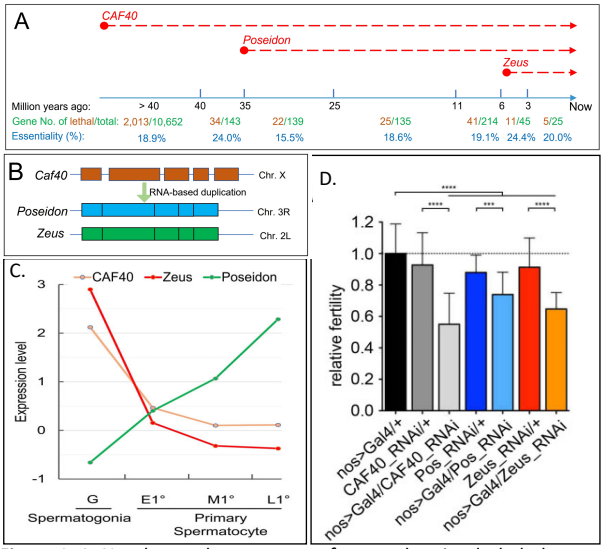

Figure 1. A. Number and percentage of genes showing lethal phenotype over total genes that originated in a period from RNAi knockdown library of *D. melanogaster*. Ages of genes were inferred based on their phylo-genetic distribution of *Drosophila* species. The gene lineages of *Caf40, Poseidon, and Zeus labeled as red dash lines [24]. B. Poseidon and Zeus* originated through RNA-based duplication (retroposition) from *Caf40*. C. *Poseidon* and Zeus show complemental expression with *Caf40* in spermatogenesis when X chromosome is inactivated in primary spermato-cyte (MXCI). D. germline-specific RNAi knockdowns detected significant reproductive effects of *Caf40, Poseidon (Pos)* and *Zeus* [25].

The carbon catabolite repression 4 (CCR4)–negative on TATA-less (NOT) complex, also known as the CCR4-NOT assembly, provides the major function of the essential deadenylase activity in eukaryotes [22]. The CCR4-NOT function is extremely conserved from yeast to humans. The *D. melanogaster* CCR4-NOT complex contains at least 8 subunits including CAF40 [19]. We recently identified two duplicates of *Caf40* in *D. melanogaster*, which were generated through RNA-based duplication (retroposition, Fig 1B) from an ancestral *Caf40* gene 3-35 million years ago [1,25]. We surprisingly found that the protein sequences of *Poseidon* and *Zeus* have evolved rapidly. Our further analyses of expression and functions of these RNA-based duplicates revealed that

they are specifically expressed during spermatogenesis with important development and male reproductive functions (Fig 1C and D), indicating that they have acquired distinct functions.

Specific Aim 1: Examine sex-specific fertility effects and cellular functions of *Poseidon*, *Zeus* and *Caf40* by creating CRISPR gene deletion lines in *D. melanogaster*.

1.1. Create gene deletion lines for *Poseidon*, *Zeus* and *Caf40*. Following our successfully implemented CRISPR-Cas9 protocols [20,24], we will generate gene deletion lines for these genes by deleting the promoters and 5' coding sequence using pairs of guide RNAs that encompass these regions. Such a design will be able to avoid potential compensation effects of frameshift mutations in duplicated genes [3,24] and possible disruption of regulatory elements for downstream genes. The deletions will be confirmed by sequencing and maintained using balancer chromosomes.

1.2. Examine the effects of gene deletions on fitness components, including sex-specific viability and fertility. Previously, our knockdown experiments showed that *Poseidon* and *Zeus* had a strong impact on fly viability and fertility. Considering the conservative nature of RNAi knockdown experiment, we infer that a knockout experiment would give us a more clear-cut result regarding the essentiality of these genes. We will use a well-developed protocol by the Long lab to measure sex-specific viability and fertility in our deletion lines [20,24]. To confirm the detected phenotypes are a result of our gene deletions, we will create rescue lines by inserting a wild-type copy back into the genomes of deletion lines.

Specific Aim 2: Detect protein interactions of POSEIDON and ZEUS. Our preliminary study [25] detected a significant binding of POSEIDON but not ZEUS to the subunit NOT1 of the CCR4-NOT complex when its X-linked parental copy *Caf40* is inactivated by the X chromosome meiosis inactivation (XCMI) during spermatogenesis. The high expression of POSEIDON at male meiosis (Fig 1C) and its interaction with CCR4-NOT suggest it might play a role in posttranscriptional and translational regulation. However, it is uncertain if POSEIDON and/or ZEUS interact with other non-NOT1 subunits of CCR4-NOT. We propose the following two lines of analyses.

2.1: Computationally predict potential interactions of POSEIDON and ZEUS with CCR4-NOT and other members of the complex and pathways.

2.1.1. Predict POSEIDON and ZEUS protein interactions with CCR4-NOT and other pathways. We will construct 3D structural models of protein-protein complexes of POSEIDON/ZEUS with NOT1 (CAF40) subunit as well as 7 other known members of the CCR4-NOT complex (CCR4, POP2, NOT1, NOT2, NOT3, NOT10 and NOT11) in *Drosophila* [19]. A recent study has shown that a CAF40 motif facilitates CCR4-NOT interaction with *Drosophila* Roquin [15]. We will take advantage of recent breakthrough in deep-learning [6] to study the structure of POSEIDON or ZEUS. We found that the POSEIDON and ZEUS structures predicted by AlphaFold2 have 90.8% of residues predicted at high-, or moderate-confidence level. We will first construct pair-wise protein-protein interacting models between POSEIDON and ZEUS with CCR4-NOT1 subunits. We will further explore whether ZEUS or POSEIDON bind with other members of the CCR4-NOT complex, including CAF1/POP2, CCR4a. We will also examine possible binding between ZEUS to potential partners identified by yeast 2-hybrid analysis, including jigr1, lms, CG11723, Rbpn-5, CG2017.

We will use AlphaFold-Multimer for this task [4]. As flexible terminal tails are subject to artificial positioning in AlphaFold-Multimer and may be in steric hindrance for protein-protein binding in the default predicted structures, we will model these terminal structures using our suite of protein loop modeling methods mDiSGRo [17,18] and PRETZEL [13]. As different domains on protein structure may become exposed, we will also experiment with modeling protein-protein interaction taking different domains from the partner proteins. We will also further carry out evolutionary dependency analysis among binding partners using tools such as GREMLIN as additional constrains for determining binding interfaces [12]. We will then integrate results using other docking methods (e.g., HADDOCK and HDOCK [21]), along with DockQ score from AlphaFold-Multimerand, to estimate the strength of binding, identify likely strong as well as transient interacting proteins, their interfaces, and key residues for binding.

2.1.2. Create expression networks to predict other interactive pathways. We will construct gene-gene interaction networks (GGIs) using a two-pronged approach. We will compute co-expression networks by statistical quantification of expressionprofile associations between pair-wise genes [26] by using Pearson's correlation coefficients across tissues for tissue RNA-seq data [2] and across cell clusters or subclusters for RNAseq data [23]. We will take genes with high correlation (e.g. $r > 0.9$) with *Poseidon* and *Zeus* as interactive candidates. Furthermore, we will infer network interactions based on 3D genome organization. We will employ the CHROMATIX method and convert existing

population Hi-C data of the loci encoding ZEUS/POSEIDON [14] and all candidate protein binding partners into large ensembles of single-cell 3D chromatin models for these at different developmental stage, as the Long & Liang labs have used before [7,16].

2.2: Experimentally detect the physical interactions of POSEIDON and ZEUS with the CCR4-NOT complex, and identify the interactomes of POSEIDON and ZEUS.

2.2.1. *Test the predicted binders from 2.1.* To experimentally verify the physical interactions of POSEIDON and ZEUS with the components of CCR4-NOT complex, we will utilize the well-established *Drosophila melanogaster* (*Dm*) S2 cell system [21]. We will first create expression constructs of HA-tagged fusions of the individual CCR4-NOT subunits that are predicted to interact with POSEIDON and ZEUS from the computational analysis as described in 2.1. We will also construct expression plasmids of the GFP-tagged POSEIDON, ZEUS, and CAF40 (as a control). The HA-tagged subunits of CCR4-NOT that are predicted to interact with POSEIDON and/or ZEUS by computational analysis from 2.1.1 will be independently co-expressed with the GFP-tagged CAF40, POSEIDON, or ZEUS. GFP-only expression plasmid will be included as a negative control. We will then carry out co-immunoprecipitation experiment using a HA antibody followed by western plot using a GFP antibody. Such co-IP experiments will confirm the predicted physical interactions between the CAF40 paralogs and the CCR4-NOT candidate components.

2.2.2. Identify the interactomes of POSEIDON and ZEUS. The findings that POSEIDON and ZEUS have quickly evolved some critically important functions in viability and fertility [24,25] suggest that the two CAF40 paralogs have established new interaction networks distinct from the one of CAF40. We propose to identify experimentally the interactomes of POSEIDON and ZEUS. The FLAG tag has been successfully used for interactomic studies in *Dm* S2 cell system. We will first create expression constructs of FLAG-tagged fusion genes of POSEIDON, ZEUS, and CAF40 (as a control) and express them independently in *Dm* S2 cells. We will then carry out co-immunoprecipitation experiments using anti-FLAG such as M2 (Sigma) and subject the isolated co-IP mixture to mass-spec analysis, using our extensive experience working with FLAG-based co-IP experiments [5]. This proposed experiment will allow us to identify interaction partners unbiasedly, including both CCR4-NOT subunits and non-CCR4-NOT binders. Comparative analysis of the obtained interactomes would also allow us to identify interaction partners that are specific to each paralog.

Nature of inter-institutional collaboration. Our three research groups will form a new collaborative team with distinctive and complementary scientific interests and expertise. Li and Long used to have a successful collaboration to investigate biochemical functions of a young chimeric gene in *Drosophila* with two existing publications. Liang will bring in exceptional computational expertise in protein structures, including machine learning analysis. Currently none of our three teams are funded for projects closely related to this proposed research. The Long group at UChicago is a leading group for the study of new genes. The Liang group of UIC is a leading group of computational biophysics and bioinformatics analyses of protein interaction. The Li group has developed outstanding experimental expertise in biochemical and molecular biology.

The criteria for measuring successes of the proposed project: identification of the interaction partners of POSEIDON and ZEUS with and/or beyond CCR4-NOT1 complex and detection of the functional consequence on reproductive processes, such as male germline functions; publication of the findings in peer-reviewed research journals and general science magazines.

Long-term funding plan beyond and timeline. We plan to apply for a joint R01 grant from NIH and a collaborative grant from NSF by using results generated from this proposed project, if funded. We expect that we will not only generate sufficient data allowing us to apply for federal grants but also open a novel and exciting new research field – origination of new protein functions. Stage 1.) April – October, 2023: Create deletion and rescue lines, Long; Computational prediction of protein interactions using AlphaFold and expression networks; Liang and Long. Stage 2.) November, 2023 – April, 2024: Experimental detection of the predicted interactions, Li lab; Genetic detection of reproductive effects, Long; Investigation of protein interactions, Liang. Stage 3.) May – November, 2024: Experimentally characterizing protein interactions, Li. Detection of any off-targeting effects by genomic sequencing and rescue experiments, Long. Comparation between predictions and experitmental confirmations, Liang and Li. Stage 4.) December, 2014 - March 2025: Data analysis and writing reports for publication: Long; Liang; and Li.

Section 3:

1. Chen SD, Ni XC, Krinsky BH, Zhang Y, Vibranovski M et al, 2012.Integration of a Nascent Gene Reshaped Global Gene Network and Impacted Sex-biased Gene Expression. *EMBO J* 31(12): 2798-2809.
2. Chen Z-X, D. Sturgill, J. Qu, H. Jiang, S. Park, N. Boley et al. 2014. Comparative validation of the D. melanogaster modENCODE transcriptome annotation. *Genome Res*. 24(7):1209–1223.
3. El-Brolosy, M.A., Z. Kontarakis, A. Rossi,, C. Kuenne, S. Gunther, N. Fukuda et al., 2019. Genetic compensation triggered by mutant mRNA degradation. *Nature* 568: p. 193-197.
4. Evans, R., M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green et al, 2021. Protein complex prediction with AlphaFold-Multimer. bioRxiv.
5. Fan, Q., K.-W. Park, Z. Du, K. A. Morano, L. Li, 2007. The Role of Sse1 in the de Novo Formation and Variant Determination of the [PSI^{+}] Prion. *Genetics*. 177(3): 1583–1593.
6. Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger et al, 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583-589.
7. Lee U, 2022. Ph.D. the University of Chicago thesis.
8. Lee YCG, Ventura IM, Rice GR, Chen DY, et al., 2019. Rapid evolution of gained essential developmental functions of a young gene via interactions with other essential genes. *Mol Biol Evol*. 36(10):2212–2226.
9. Lewin, B. 2007. Genes IX. Jones & Bartlett Learning.
10. Long M, VanKuren N, Chen SD, Vibranovski M. 2013. New gene evolution: little did we know. *Annual Review of Genetics* 47:325-351.
11. Kasinathan B, ColmenaresSU 3rd, McConnell H,Young JM,Karpen GH, Malik HS. 2020. Innovation of heterochromatin functions drives rapid evolution of essential ZAD-ZNF genes in Drosophila. *eLife* 9:e63368.
12. Ovchinnikov, S., H. Kamisetty, D. Baker, 2014. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife* 3: e02030.
13. Perez-Rathke,A., M.A. Fahie, C. Chisholm, J. Liang & M. Chen, 2018. Mechanism of OmpG pH-Dependent Gating from Loop Ensemble and Single Channel Studies. *J Amer Chem Soc* 140 (3), 1105-1115.
14. Perez-Rathke, A., Q. Sun, B. Wang *et al,* 2020. CHROMATIX: computing the functional landscape of many-body chromatin interactions in transcriptionally active loci from deconvolved single cells. *Genome Biol* 21, 13.
15. Sgromo, A., T. Raisch, P. Bawankar, D. Bhandari, Y. Chen, et al. 2017. A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin. Nat Comm. 8:14307–14316.
16. Sun, I., A. Perez-Rathke, D. M. Czajkowsky, Z. Shao, and J. Liang, 2021. High-resolution single-cell 3D-models of chromatin ensembles during Drosophila embryogenesis. *Nature Communications* 12*,* 205.
17. Tang, K., W.K.S. Wong, J. Liu, J. Zhang, and J. Liang, 2015. Conformational sampling and structure prediction of multiple interacting loops in soluble andβ-barrel membrane proteins using multi-loop distance-guided chain-growth Monte Carlo method. *Bioinformatics* 31(16), 2015, 2646–2652 .
18. Tang, K., J. Zhang, and J. Liang, 2017. Distance-Guided Forward and Backward Chain-Growth Monte Carlo Method for Conformational Sampling and Structural Prediction of Antibody CDR-H3 Loops. *J. Chem. Theory Comput*. 13: 380-388.
19. Temme, C., M. Simonelig & E. Wahle, 2014. Deadenylation of mRNA by the CCR4–NOT complex in *Drosophila*: molecular and developmental aspects. *Front Genet*. 5, 143.
20. VanKuren NW, Long MY. 2018. Gene duplicates resolving sexual conflict rapidly evolved essential gametogenesis functions. *Nat Ecol Evol*. 2(4):705–712.
21. van Zundert, G.C.P., J.P.G.L.M. Rodrigues, M. Trellet, et al, 2016. The HADDOCK2.2 webserver: User-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.*, 428, 720-725.
22. Wahle E, Winkler GS. 2013. RNA decay machines: deadenylation by the CCR4–NOT and Pan2-Pan3 complexes. *Biochim Biophys Acta*. 1829(6-7):561–570.
23. Witt E, Benjamin S, Svetec N, Zhao L. Testis single-cell RNA-seq reveals the dynamics of de novo gene transcription and germline mutational bias in Drosophila. *Elife*. 2019; 8:e47138.
24. Xia, S., N.W.VanKuren, C. Chen, L. Zhang, et al, 2021. Genomic Analyses of New Genes and Their Phenotypic Effects Reveal Rapid Evolution of Essential Functions in Drosophila Development. *PLoS Genetics*. **17**(7): e1009654.
25. Xia, S., I.M. Ventura, A. Blaha, A. Sgromo, S. Han, E. Izaurralde, & M. Long, 2022. Rapid Gene Evolution in an Ancient Post-transcriptional and Translational Regulatory System Compensates for Meiotic X Chromosomal Inactivation. *Mol. Biol. Evol.* 39(1): msab296.
26. Zhang, W., P. Landback, A.R. Gschwend, B. Shen, M. Long, 2015. New genes drive the evolution of gene interaction networks in the human and mouse genomes. *Genome Biol*, 2015. 16: p. 202.

Section 4:

The University of Chicago Investigator's Budget Form

Program Director/Principal Investigator (Last, First, Middle): Manyuan Long

DETAILED BUDGET FOR INITIAL BUDGET PERIOD
DIRECT COSTS ONLY

FROM 10/01/20 THROUGH 09/30/22

List PERSONNEL *(Applicant organization only)*
Use Cal, Acad, or Summer to Enter Months Devoted to Project
Enter Dollar Amounts Requested *(omit cents)* for Salary Requested and Fringe Benefits

NAME	ROLE ON PROJECT	Cal. Mnths	Acad. Mnths	Summer Mnths	INST.BASE SALARY	SALARY REQUESTED	FRINGE BENEFITS	TOTAL
Manyuan Long	PD/PI	0.20			243,804	4,063	1,028	5,091
Postdoc	RA	6			54,840	27,420	6,937.	34,357
SUBTOTALS →						31,483	7,965.	39,448

CONSULTANT COSTS

EQUIPMENT *(Itemize)*

SUPPLIES *(Itemize by category)*

Food for flies: 700
Lab supplies: 500
Supplies subtotal: 1,200

TRAVEL

INPATIENT CARE COSTS

OUTPATIENT CARE COSTS

ALTERATIONS AND RENOVATIONS *(Itemize by category)*

OTHER EXPENSES *(Itemize by category)*

Publication

CONSORTIUM/CONTRACTUAL COSTS | DIRECT COSTS

SUBTOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD *(Item 7a, Face Page)* $ 40,648

CONSORTIUM/CONTRACTUAL COSTS | FACILITIES AND ADMINISTRATIVE COSTS

TOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD $ 40,648

PHS 398 (Rev. 08/12 Approved Through 8/31/2015) Page ____ OMB No. 0925-0001
Form Page 4

The University of Chicago Investigator's Budget Form

Program Director/Principal Investigator (Last, First, Middle): Manyuan Long

BUDGET FOR ENTIRE PROPOSED PROJECT PERIOD
DIRECT COSTS ONLY

BUDGET CATEGORY TOTALS	INITIAL BUDGET PERIOD *(from Form Page 4)*	2nd ADDITIONAL YEAR OF SUPPORT REQUESTED	3rd ADDITIONAL YEAR OF SUPPORT REQUESTED	4th ADDITIONAL YEAR OF SUPPORT REQUESTED	5th ADDITIONAL YEAR OF SUPPORT REQUESTED
PERSONNEL: *Salary and fringe benefits. Applicant organization only.*	39,448	40,631			
CONSULTANT COSTS					
EQUIPMENT					
SUPPLIES	1,200	1,100			
TRAVEL					
INPATIENT CARE COSTS					
OUTPATIENT CARE COSTS					
ALTERATIONS AND RENOVATIONS					
OTHER EXPENSES		1,021			
DIRECT CONSORTIUM/ CONTRACTUAL COSTS					
SUBTOTAL DIRECT COSTS *(Sum = Item 8a, Face Page)*					
F&A CONSORTIUM/ CONTRACTUAL COSTS					
TOTAL DIRECT COSTS	40,648	42,752			

TOTAL DIRECT COSTS FOR ENTIRE PROPOSED PROJECT PERIOD	83,400

JUSTIFICATION. Follow the budget justification instructions exactly. Use continuation pages as needed.

UIC Investigator's Budget Form

Program Director/Principal Investigator (Last, First, Middle): Liang, Jie

DETAILED BUDGET FOR INITIAL BUDGET PERIOD
DIRECT COSTS ONLY

FROM 03/01/2023 THROUGH 02/28/2024

List PERSONNEL *(Applicant organization only)*
Use Cal, Acad, or Summer to Enter Months Devoted to Project
Enter Dollar Amounts Requested *(omit cents)* for Salary Requested and Fringe Benefits

NAME	ROLE ON PROJECT	Cal. Mnths	Acad. Mnths	Summer Mnths	INST.BASE SALARY	SALARY REQUESTED	FRINGE BENEFITS	TOTAL
Liang Jie	PD/PI			0.24	257,470	5,149	1,830	6,979
Research Assistant	RA		4.5	1.2		27,108	1,342	28,450
SUBTOTALS →						32,257	3,172	35,429

CONSULTANT COSTS	
EQUIPMENT *(Itemize)*	
SUPPLIES *(Itemize by category)*	2,641
TRAVEL	3,000
INPATIENT CARE COSTS	
OUTPATIENT CARE COSTS	
ALTERATIONS AND RENOVATIONS *(Itemize by category)*	
OTHER EXPENSES *(Itemize by category)*	

CONSORTIUM/CONTRACTUAL COSTS	DIRECT COSTS	
SUBTOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD *(Item 7a, Face Page)*		$ 41,070
CONSORTIUM/CONTRACTUAL COSTS	FACILITIES AND ADMINISTRATIVE COSTS	
TOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD		$ 41,070

UIC Investigator's Budget Form

Program Director/Principal Investigator (Last, First, Middle): Liang, Jie

BUDGET FOR ENTIRE PROPOSED PROJECT PERIOD
DIRECT COSTS ONLY

BUDGET CATEGORY TOTALS	INITIAL BUDGET PERIOD *(from Form Page 4)*	2nd ADDITIONAL YEAR OF SUPPORT REQUESTED	3rd ADDITIONAL YEAR OF SUPPORT REQUESTED	4th ADDITIONAL YEAR OF SUPPORT REQUESTED	5th ADDITIONAL YEAR OF SUPPORT REQUESTED
PERSONNEL: *Salary and fringe benefits. Applicant organization only.*	35,429	36,590			
CONSULTANT COSTS					
EQUIPMENT					
SUPPLIES	2,641	2,640			
TRAVEL	3,000	3,000			
INPATIENT CARE COSTS					
OUTPATIENT CARE COSTS					
ALTERATIONS AND RENOVATIONS					
OTHER EXPENSES					
DIRECT CONSORTIUM/ CONTRACTUAL COSTS					
SUBTOTAL DIRECT COSTS *(Sum = Item 8a, Face Page)*	41,070	42,230			
F&A CONSORTIUM/ CONTRACTUAL COSTS					
TOTAL DIRECT COSTS	41,070	42,230			
TOTAL DIRECT COSTS FOR ENTIRE PROPOSED PROJECT PERIOD					$ 83,300

JUSTIFICATION. Follow the budget justification instructions exactly. Use continuation pages as needed.

see attached

Northwestern University Investigator's Budget Form

Program Director/Principal Investigator (Last, First, Middle): Liming Li

DETAILED BUDGET FOR INITIAL BUDGET PERIOD DIRECT COSTS ONLY	FROM 03/01/2023	THROUGH 02/29/25

List PERSONNEL *(Applicant organization only)*
Use Cal, Acad, or Summer to Enter Months Devoted to Project
Enter Dollar Amounts Requested *(omit cents)* for Salary Requested and Fringe Benefits

NAME	ROLE ON PROJECT	Cal. Mnths	Acad. Mnths	Summer Mnths	INST.BASE SALARY	SALARY REQUESTED	FRINGE BENEFITS	TOTAL
Liming Li	PD/PI	0.24			128,125	2,562	715	3,277
TBD	Technician	6			40,000	21,000	5,580	26,580
								0
								0
								0
								0
								0
SUBTOTALS →						23,562	6,295	29,857

CONSULTANT COSTS	
EQUIPMENT *(Itemize)*	
SUPPLIES *(Itemize by category)* core facility, antibodies, media, reaction kits, chemicals...	15,426
TRAVEL	1,000
INPATIENT CARE COSTS	
OUTPATIENT CARE COSTS	
ALTERATIONS AND RENOVATIONS *(Itemize by category)*	
OTHER EXPENSES *(Itemize by category)* Publication - $500	500
CONSORTIUM/CONTRACTUAL COSTS	DIRECT COSTS
SUBTOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD *(Item 7a, Face Page)*	$ 46,783
CONSORTIUM/CONTRACTUAL COSTS	FACILITIES AND ADMINISTRATIVE COSTS
TOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD	$ 46,783

Northwestern University Investigator's Budget Form

Program Director/Principal Investigator (Last, First, Middle): Liming Li

BUDGET FOR ENTIRE PROPOSED PROJECT PERIOD
DIRECT COSTS ONLY

BUDGET CATEGORY TOTALS	INITIAL BUDGET PERIOD *(from Form Page 4)*	2nd ADDITIONAL YEAR OF SUPPORT REQUESTED	3rd ADDITIONAL YEAR OF SUPPORT REQUESTED	4th ADDITIONAL YEAR OF SUPPORT REQUESTED	5th ADDITIONAL YEAR OF SUPPORT REQUESTED
PERSONNEL: *Salary and fringe benefits. Applicant organization only.*	29,857	26,971			
CONSULTANT COSTS					
EQUIPMENT					
SUPPLIES	15,426	9,246			
TRAVEL	1,000				
INPATIENT CARE COSTS					
OUTPATIENT CARE COSTS					
ALTERATIONS AND RENOVATIONS					
OTHER EXPENSES	500				
DIRECT CONSORTIUM/ CONTRACTUAL COSTS					
SUBTOTAL DIRECT COSTS *(Sum = Item 8a, Face Page)*	46,783	36,217	0	0	0
F&A CONSORTIUM/ CONTRACTUAL COSTS					
TOTAL DIRECT COSTS	46,783	36,217	0	0	0
TOTAL DIRECT COSTS FOR ENTIRE PROPOSED PROJECT PERIOD					**$** 83,000

JUSTIFICATION. Follow the budget justification instructions exactly. Use continuation pages as needed.

Personnel

Senior/Key Personnel

Budget Justification

Dr. Manyuan Long, Principal Investigator, will devote 0.2 calendar months of effort to this project. We are requesting 0.2 calendar months of salary support per year for Dr. Long. Dr. Long will design, supervise and organize the entire scientific experiments for specific aim 1 and computational analyses in details and he will also lead the writing of research reports.

All academic appointments in the Biological Science Division are based on twelve months and effort on projects is spread over the calendar year. This request for CBC salary support is, we believe, consistent with the CBC limit of 2% of regular academic year salary for investigators with academic-year appointments, and also reflects the policy of the Division of the Biological Sciences and the Pritzker School of Medicine that faculty recover appropriate salary support from grants and contracts proportionate to the effort devoted to the project.

Other Personnel

Postdoctoral Scholar (to be hired) will devote 12 calendar months of effort annually to this project. This postdoctoral scholar will be in charge of creating deletion lines with rescue lines and testing reproductive effects of two new genes. The postdoctoral scholar is expected to construct computationally generate co-expression networks of new genes with all other related genes. The scholar will also collaborate the group of Li to study gene interaactions. We request 12 calendar months of salary for this postdoctoral scholar.

Fringe Benefits

The University of Chicago charges a fringe benefits rate of 25.3% on faculty and postdoctoral scholar salaries. The University of Chicago does not charge fringe benefits for Graduate Research Associates.

Supplies

The cost for flies foods and chemicals in preparing CRISPR deletions.

Travel

Publication

The other expenses are used to pay a part of publication of the scientific results obtained in this project.

BUDGET JUSTIFICATION

Personnel

Senior/Key Personnel
Dr. Jie Liang, will devote 2% of efforts to this project. Dr. Liang will design, supervise, and organize the computational studies in predicting protein-binding partners of ZEUS and POSEIDON, and will work with Dr. Long and Dr. Li in analyzing overall protein-protein interaction networks.

Other Personnel
Bowei Ye, PhD student, will devote 9+1.8 months as research assistant at the level of 50% efforts. He will carry out will develop necessary code and scripts, and carry out simulations. He will also run simulations related to terminal regions/loop remodeling, and will work on porting our PRETZEL and mDisGro code so they can be run at UIC High Performance Computing cluster. He will also carry out detailed analysis.

Fringe Benefits
The institutional fringe rate for senior personnel is 35.54% for Year 1, 3.64%/11.52% for Graduate students' academic year/summer for Year 1. There is 3% of increase for Year 2

Travel
Liang and Ye will travel once a year to conference to present progress and results from this project to scientists in the field. $1,500+$1,500/year are requested.

Materials and Supply
A total of $5,281 is requested for supplies. This will be used to purchase basic supplies such as a workstation, monitors, hard drives, printers, paper and other necessities for PI and PhD student.

Budget Justification

Personnel

Senior/Key Personnel

Dr. Liming Li, Principal Investigator, will devote 0.24 calendar months of effort to this project. We are requesting 0.24 calendar months of salary support in two years for Dr. Li. She will design, supervise and organize the entire scientific experiments for Specific Aims 2.2 to identify the physical interaction partners of ZEUS and POSEIDON. She will work with Drs. Long and Liang on analyzing results and writing scientific reports of the proposed project.

All academic appointments in the Feinberg Medical School of Northwestern University are based on
twelve months and effort on projects is spread over the calendar year. This request for CBC salary support is consistent with the CBC limit of 2% of regular academic year salary for investigators with academic year appointments, and also reflects the policy of Northwestern University that faculty recover appropriate salary support from grants and contracts proportionate to the effort devoted to the project.

Other Personnel

To be named, research technician assistant, will devote 6 calendar months of effort to this project for two years. This technician will be in charge of carrying out the experimental plans made by the P. I., including all related to specific aim 2.2 and collaboration with the personnel from other two research teams in the experiments related to specific aim 2.

Employee benefits have been calculated based on the following rates:

- *09/01/22 – 08/31/23…. 27.9%*
- *09/01/23 – 08/31/24 and thereafter…. 27.9% (provisional)*

Travel

Liming Li plans to travel once for discussing about the scientific results from this project with scientists in related fields, with $1,000 provided by CBC.

Publication

$500 provided by CBC is planned to use to pay publication of the scientific results obtained in this project.

Services

The Northwestern University Proteomics Core Facility will be used to analysis the interactomes of ZEUS and POSEIDON as described in specific aim 2.

Section 5:

OMB No. 0925-0001 and 0925-0002 (Rev. 09/17 Approved Through 03/31/2020)

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: Long, Manyuan

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

POSITION TITLE: Edna K. Papazian Distinguished Service Professor, the University of Chicago

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
Sichuan Agricultural University, China	M.S./B.S.	02/1985	Genetics & Plant Breeding
University of California, Davis, CA	Ph.D./M.S.	12/1992	Genetics
Harvard University, MA	Postdoctoral F.	10/1997	Genetics

A. Personal Statement

My research career until so far has been focused on a scientific problem I defined and started when I was a graduate student at UC Davis in the early 1990s: How do evolutionary new genes originate and evolve? Since I reported the first ever known new gene in African flies in my 1992 doctoral thesis, I have investigated several major scientific issues regarding the new gene problem: rates of new gene origination, patterns of new gene evolution, molecular mechanisms that generate new genes, evolutionary forces acting on new genes, phenotypic effects and functions of new genes, genomic evolution impacted by new genes. These researches have helped shape new chapters in textbooks of evolution, genetics and genomics. Consequently, I also trained dozens of graduate students and postdocs who are principal investigators in research institutes and universities. Today, evolutionary new genes have become important topics in basic molecular biology and medical research community.

Since 2010, we have explored functions and phenotypic effects of new genes and underlying molecular mechanisms. Some of previously unexpected observations that have not been predicted by conventional theories included rapid evolution of important gene functions and generation of new protein coding functions out of noncoding DNA sequences. The new data we collected from experiments of molecular genetics, while challenging old theories, raised new and important scientific problems and made changes in the ways scientists view life and practically improve our health conditions. For example, our discovery of the mutational process to generate new genes through *de novo* origination was hailed by the Medscape journal as one of the "15 Studies that Challenged Medical Dogma in 2019". In this proposal, we further described how new proteins created from gene duplication changed a most conserved functional system that governs posttranscriptional and translational regulation in all eukaryotes.

1. Chen SD, Ni XC, Krinsky BH, Zhang Y, Vibranovski M, White KP and Long M, 2012. Integration of a Nascent Gene Reshaped Global Gene Network and Impacted Sex-biased Gene Expression. *EMBO J* 31(12): 2798-2809.

2. Xia S, VanKuren NW, Chen C, Zhang L, Kemkemer C, Shao Y, Jia H, Lee U, Advani S, Gschwend A, Vibranovski M, Chen S, Zhang YE and Long M, 2021. Genomic Analyses of New Genes and Their Phenotypic Effects Reveal Rapid Evolution of Essential Functions in Drosophila Development. *PLoS Genet*. 17(7): e1009654.
3. Xia S, Ventura I, Blaha A, Sgromo A, Han S, Izaurralde E, Long M, 2022. Rapid gene evolution in an ancient post-transcriptional and translational regulatory system compensates for meiotic X chromosomal inactivation. *Mol Biol Evol* 39(1): msab296.
4. Krinsky BH, Arthur RK, Xia S, Sosa D, Arsala D, White KP, Long M, 2022. Rapid Cis-Trans Coevolution Driven by a Novel Gene Retroposed from a Eukaryotic Conserved CCR4–NOT Component in *Drosophila*. *Genes* 13(1), 57.

B. Positions and Honors

Positions and Employment

Edna K. Papazian Distinguished Service Professor, Department of Ecology and Evolution and the College, The University of Chicago; a faculty member of the Committees on Genetics and Evolutionary Biology, The University of Chicago. Started July 1, 2011, Permanent appointment.

Full professor with tenure, Department of Ecology and Evolution and The College, The University of Chicago; a faculty member of the Committees on Genetics and Evolutionary Biology, The University of Chicago. January 2005 – June 2011.

Senior Fellow, The Institute for Genomics & Systems Biology, The University of Chicago and Argonne National Laboratory, April 2007 – current.

Associate Professor with tenure, Department of Ecology and Evolution and The College, The University of Chicago; a faculty member of the Committees on Genetics and Evolutionary Biology, The University of Chicago. July 1, 2003 – December 31, 2004.

Assistant professor, Department of Ecology and Evolution and the College; A faculty member of the Committees on Genetics and Evolutionary Biology, The University of Chicago. November 1, 1997 – June 30, 2003.

Honors:

2022 John Simon Guggenheim Memorial Fellowship for Biology, the Guggenheim Foundation (One of 3 awardees in biology, one of 23 awardees in natural sciences and one of 180 awardees from 51 fields in sciences, humanity and arts in U.S. and Canada).
2022 Ray Wu Award, the Ray Wu Society (the Chinese Biological Investigator Society in U.S.) (one of 2 awardees).
2020 Distinguished Investigator Award, Senior Award, the University of Chicago, Division of Biological Sciences (one of 2 awardees).
2019 15 Studies that Challenged Medical Dogma in 2019, Medscape.
2014 AAAS Fellow, the American Association for the Advancement of Science.
2013 Honorary Visiting Professor, Division of Theoretical Genetics and Evolution, National Institute
2016 of Genetics, Japan
2011 Inaugural Edna K Papazian Distinguished Service Professor, the University of Chicago
2003 CAREER Award, National Science Foundation
1999 David & Lucile Packard Fellowship for Science & Engineering, the Davis & Lucile Packard Foundation
1993 Allen Marr Prize, the University of California, Davis (the only awardee).

C. Contributions to Science

(i) Making the discovery of first ever known evolutionary new gene and proposing new gene evolution as a new topic in study of genes and evolution.

Genes have been often viewed conserved in evolution. Evolutionary new genes had been thought unlikely until the early 1990s. In my doctoral thesis research, I made a surprising finding from my population genetic study of two closely related African Drosophila species: a new gene named Jingwei was generated by retroposition of a gene encoding alcohol dehydrogenase and insertion into another unrelated gene in a most recent common ancestor 3 million years ago. This gene provided the first insight into how a new gene originated in genetic and evolutionary mechanisms. For the first time, we proposed a new concept of evolutionary new gene in the paper we published in Science. We furthermore demonstrated that the gene evolved a new molecular function of recruitment hormone and its exon-intron structure is evolving in a transition of polymorphism. This case study has been written into a few textbooks of evolution and genetics.

1. Long, M., C. H. Langley, 1993. Natural selection and the origin of *jingwei*, a chimeric processed functional gene in *Drosophila*. *Science* 260: 91-95.
2. Zhang, J., A. M. Dean, F. Brunet and M. Long 2004. Evolving functional diversity in new genes of *Drosophila*. *Proc Natl Acad Sci USA.* 101: 16246 -16250.
3. Llopart, A., J. M. Comeron, F. G. Brunet, D. Lachaise, M. Long. 2002. Intron presence / absence polymorphism in Drosophila driven by positive Darwinian selection. *Proc Natl Acad Sci USA* 99: 8121-8126.
4. Zhang J, Long M, and Li L, 2005. Translational effects of differential codon usage among intragenic domains of the young Drosophila gene. *Biochim Biophys Acta* 1728:135-142.

(ii) Demonstrating that evolutionary new genes are common, showing variable rates and regular patterns in various organisms.

(iii)

Soon after we published the Jingwei gene and its parental gene in the early 2020s, we investigated the generality of new gene evolution. We experimentally designed genomic screening across Drosophila species and other organisms and computationally identified genes that evolved recently using genome sequences in the species ranging from Drosophila to vertebrates, especially human and related primate species. We also examined the exon-intron structures across eukaryotes and detected an extensive exon shuffling that yielded chimeric proteins. These works revealed that *Jingwei* is not an exception but new gene evolution is common in organisms. Furthermore, we identified a gene traffic between sex chromosomes and autosomes, which is unidirectional in Drosophila and bi-directional in mammals.

5. Betrán, E., K. Thornton, and M. Long 2002. Retroposed new genes out of the *X* in *Drosophila. Genome Research.* 12: 1854-1859.
6. Emerson J.J., H. Kaesmann, E. Betrán and M. Long 2004. Extensive gene traffic on the human X chromosome. *Science* 303: 537-540.
7. Zhang YE, Landback P, Vibranovski MD, Long M, 2011. Accelerated recruitment of new brain development genes into the human genome. *PLoS Biology* 9(10): e1001179..
8. Vibranovski MD, Zhang Y, Long M, 2009. General gene movement off the X chromosome in the *Drosophila* genus. *Genome Research* 19: 897-903.

(iv) Revealing molecular mechanisms that create new genes.

The classic model for new gene evolution was the gene duplication proposed by Hermann Muller in 1936. We investigated this model and other models for the mechanisms involved to generate initial structure of a new gene. We showed that retroposition, exon shuffling and gene recombination intermediated by

repetitive elements occur frequently, often in conjunction with gene duplication. Today, we have known at least a dozen of mechanisms involved in the generation of new gene structure.

9. Tan S, Cardoso-Moreira M, Shi W, Zhang D, Huang J, Mao Y, Jia H, Zhang Y, Liu Z, Huang X, Long M, Zhang YE, 2016. LTR-mediated retroposition as a mechanism of RNA-based duplication in metazoans. *Genome Research* 26:1663–1675 .
10. Yang S., Arguello R, Li X, Ding Y, Zhou Q, Chen Y, Zhang Y, Zhao R, Brunet F, Peng L, Long M, Wang W., 2008. Repetitive Elements-mediated Recombination as a Mechanism for New Gene Origination in *Drosophila. PLoS Genetics* 4(1): e3. 0001-0010.
11. Long, M., C. Rosenberg, W. Gilbert 1995. Intron phase correlations and the evolutiointron/exon structure of genes. *Proc Natl Acad Sci USA* 92: 12495-12499.
12. Long M, VanKuren N, Chen SD, Vibranovski M, 2013. New Gene Evolution: Little Did We Know. *Annual Review of Genetics* 47: 325-351.

(v) Detection of evolutionary forces on the new genes and genomes.

Using population genetic analyses and substitution analyses, we detected positive selection plays a common role in new gene evolution, consequently driving rapid changes of new genes in gene structure and sequences (20). Particularly, we detected strong selection on the polymorphic duplicates, the first stage in new gene evolution, in Drosophila populations. Recently, we were surprised by the strong sexually antagonistic selection that drove evolution a pair of young duplicates that have been fixed in D. melanogaster populations (<0.2 million years old). We found nonadaptive forces such as sexual conflict and sexual selection are often in determination of evolutionary fates of new gene mutants.

13. VanKuren NW and Long M, 2018. Gene duplicates resolving sexual conflict rapidly evolved essential gametogenesis functions. *Nature Ecology & Evolution* 2(4):705-712.
14. Emerson JJ, Cardoso-Moreira M, Borevitz JO, Long M, 2008. Natural selection shapes genome wide patterns of copy number polymorphism in D. melanogaster. *Science* 320:1629-1631.
15. Wang, W., K. Thornton, A. Berry, and M. Long. 2002. Nucleotide variation along the *Drosophila melanogaster* fourth chromosome. *Science* 295:134-137.
16. Arguello JR, Chen Y, Yang S, Wang W, and Long M, 2006. An X-linked Testes Chimeric Gene by Illegitimate Recombination in *Drosophila. PLoS Genetics* 2: 0745-0754. e77.

(vi) Phenotypic and functional studies of new genes that revealed unexpected effects.

The rapid evolution with various types of natural selection, as the aforementioned studies, suggest functional and phenotypic effects new genes bring about. We thus carried out functional and phenotypic studies of new genes and detected significant effects in development, behaviors, brains and reproductive traits. Remarkably, we found that the directional gene traffic from the X to autosomes was associated with the meiotic X chromosome inactivation in male germline cell in Drosophila. Surprisingly, we found that the new genes in Drosophila evolved essential functions quickly, suggesting a rapid evolutionary genetic basis underlying development process. These conceptual novel results we obtained 10 years ago were recently confirmed by independent labs, standing the test of time.

17. Dai H, Chen Y, Chen S, Mao Q, Kennedy D, Landback P, Eyre-Walker A, Du D, Long M, 2008. The evolution of courtship behaviors through the origination of a new gene in *Drosophila. Proc Natl Acad Sci USA*. 105:7478-83.
18. Vibranovski MD, Lopes HF, Karr TL, Long M, 2009. Stage-specific expression profiling of Drosophila spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis expressed genes. *PLoS Genet* 5(11): e1000731.

19. Chen S, Zhang Y, and Long M, 2010. New genes in Drosophila quickly become essential. *Science* 330: 1682-1685.
20. Chen SD, Ni XC, Krinsky BH, Zhang Y, Vibranovski M, White KP and Long M, 2012. Integration of a Nascent Gene Reshaped Global Gene Network and Impacted Sex-biased Gene Expression. *EMBO J* 31(12): 2798-2809.

(vii) Origin of new genes from scratch.

The possibility that a gene is created de novo from previously noncoding sequences that would translate a protein with likely random amino acid sequences was rejected by founders of molecular biology and theoretical architect in evolutionary biology, e.g. Francois Jacob, Jacque Monod, and Ernst Mayr. The human-specific functions were viewed only from a process of duplication with limited modification. We show that these views are not what data and revealed processes told.

21. Zhang L, Ren Y, Yang T, Li G, Chen J, Gschwend AR, Yu Y, Hou G, Zi J, Zhou R, Wen B, Zhang J, Chougule K, Wang M, Copetti D, Peng Z, Zhang C, Zhang Y, Ouyang Y, Wing RA, Liu S, Long M, 2019. Rapid evolution of protein diversity by *de novo* origination in *Oryza*. *Nature Ecology & Evolution*. 3(4), 679–690. Highlighted with a News & Views article by Bornberg-Bauer E and Heames B (2019. Becoming a de novo gene. *Nature Ecology & Evolution* 3(4), 524–525) and a News Feature Article by Adam Levy (2019. Genes from the junkyard. *Nature* 574, 314-316.) Selected into "15 Studies That Challenged Medical Dogma in 2019" in the *Medscape* by Topol EJ.
22. Mortola E and Long M, 2021. Turning Junk into Us: How Genes Are Born. *American Scientist*. 109(5/6): 174-181.
23. Chen S, Krinsky B, Long M, 2013. Chen SD, Krinsky B, Long M, 2013. New genes as drivers of phenotypic evolution. *Nature Reviews Genetics* 14 (9): 645-660.
24. Zhang Y, Vibranovski MD, Landback P, Long M, 2010. Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on X chromosome. *PLoS Biology* 8(10): e1000494.

D. Additional Information: Research Support

Ongoing Research Support

National Science Foundation (2020667) 09/08/2020~09/07/2023

*"*Test the Molecular Mechanism of New Gene Evolution under Sexual Conflict in Drosophila*"* (Long, PI)

This project aims to explore ancient phenotypes in the most rest common ancestor of two tandem duplicates that evolved opposite functions in spermatogenesis and oogenesis in Drosophila. The evolutionary process to the extant Apollo and Artemis genes under sexually antagonistic selection is expected to understand.

National Institutes of Health (R01GM116113-01A1) 09/01/2016 – 08/31/2023

*"*Reference-quality Drosophila genome assemblies for evolutionary analysis of previously in accessible genomic regions*"* (Long, multiple PI)

The goal of this program is to create high-quality genomes of 12 Drosophila species for evolutionary analysis, using third generation sequencing (PacBio) in combination with bacterial artificial chromosomes libraries. The Long team is expected to investigate new gene evolution using these genome assemblies.

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: Liang, Jie

eRA COMMONS USER NAME (credential, e.g., agency login): jliang@uic.edu

POSITION TITLE: Richard and Loan Hill Professor of Bioengineering

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
Fudan University, Shanghai, China	BS	07/1986	Biophysics
University of Illinois at Urbana-Champaign	PhD	05/1994	Biophysics
University of Illinois at Urbana-Champaign	MCS	05/1995	Computer Science
Computer Science/Beckman Institute/Ntnl Ctr for Supercomputing Apps, UIUC	NSF Postdoc	1994-1996	Biophysics/Biogeometry
NSF Institute of Math Appl / U of Minnesota	Postdoc	1996-1997	Math and Biochemistry

A. Personal Statement

Research in my lab has been focused on constructing quantitative models and developing computational methods to study complex biological systems. I have strong expertise in algorithm development and modeling, with significant experience in analysis of structures of biopolymers (chromatins, proteins, and RNAs).

In **protein biophysics**, my lab has developed state-of-the-art tools for computing protein binding pockets and voids, as well as general metric measurements based on the alpha shapes and dual complexes, an early application of persistent homology and topological data analysis. The CastP server developed in my lab is widely used by structural biologists. In **3D chromatin folding**, my lab developed the Constrained Self-Avoiding Chromosome (C-SAC) model, the geometric sequential importance sampling technique, fractal Monte Carlo, and the CHROMATIX method (Gursoy et al, 2014; Gursoy et al, 2917a, 2017b; Perez-Rathke et al, 2020). Our work with Dr. Amy Kenter elucidated the physical basis of the scaling properties of chromosomes observed in FISH and Hi-C (Gursoy et al 2014). With a few simple biological landmarks, our model can reproduce genome-wide Hi-C measurements of budding yeast, including both inter- and intra-chromsome interactions (Gursoy et al 2017a). In another collaboration with Dr. Kenter, our 3D ensemble model also predicted functional interactions and uncovered a novel site of 3-body interactions in the alpha-globin locus (Gursoy et al 2017b).

Recently, my laboratory has developed the CHROMATIX algorithms for reconstruction of ensemble 3D structures of single-cell chromosomes from population Hi-C and for identifying many-body chromatin interactions (Perez-Rathke et al, 2020; Sun et al, 2021). Our results revealed that a small set of specific interactions (5–6% in Drosophila) of measured Hi-C can fold chromatin. Without any adjustable parameter, our model can fold chromatin chains into 3D achieving high accuracy (R=0.96-0.97). Our single-cell 3D chromatin conformations are in excellent agreement with single-cell experimental measurements, 3D-FISH measurements, and can quantitatively characterize

chromatin heterogeneity (Sun et al, 2021). Overcoming the limitation of the pairwise and population-averaged nature of Hi-C data, we have uncovered a large number of significant many-body interactions, and constructed detailed functional landscape of higher-order many-body interactions for all 39 highly-active transcriptional loci with ≥2 super-enhancers bounded by a TAD in human lymphocytes. We also showed that epigenetic marks can predict regions of many-body interactions (Perez-Rathke et al, 2020).

There has been significant further development of prototypes of algorithms and software described **in two recent publications** (Perez-Rathke, *Genome Biology*, 2020; Sun et al, *Nature Communications*, 2021).

Overall, my lab has a long history of developing effective deep sampling algorithms of chain polymers, including constrained chain-growth sequential Monte Carlo method for generating protein polymer conformations from contact maps, pseudoknotted RNA structure predictions, protein loop ensemble generation and structure prediction, as well as algorithms for sequence order independent protein structure alignment. In **computational cell and tissue modeling,** my lab developed the dynamic cell finite element method (dyCelFEM), the first method that can simulate thousands of cells with complex cell shapes in large scale migration under both biochemical and mechanical directional cues. Other areas where work from my lab has made significant contributions include **membrane protein design** and **stochastic reaction kinetic and cellular network**, which are described in Contributions to Science.

The Liang lab has **a strong track record** in disseminating computational tools and techniques. The CastP webserver (sts.bioe.uic.edu) is widely used by the structural biology community. It is based on alpha shape and persistent homology and processes topological and geometric computation on all Protein Data Bank structures, with results organized in a database and integrated with other genomic information. Our publication of Castp 3.0 in 2018 has been cited >680 times. Cumulatively, this server has been cited 3,270 times since its inception. Additional tools include TopoSNP for nsSNP analysis, LIPS for membrane protein lipid face prediction, TMBB-EXPLORER for membrane protein structure prediction, DiSGro tools for protein loop modeling, and ACME for stochastic network.

As of March 2022, I have advised 21 PhD students who graduated. I have authored >160 publications, which attracted >12,300 citations.

Ongoing and recently completed projects that I would like to highlight include:

R03OD032628 (Liang: PI) 9/22/2021 - 9/21/2022
NIH/Common Fund
Constructing High-Resolution Ensemble Models of 3D Single-Cell Chromatin Conformations of eQTL Loci from Integrated Analysis of 4DN-GTEx Data towards Structural Basis of Differential Gene Expression

R21AI126308-01A1 (Liang: PI, Kenter: Co-I) 8/18/2017 – 7/31/2019
NIH/NIAID
Constructing Ensembles of 3D Structures of Igh Locus and Predicting Novel Chromosomal Interactions

R35GM127084-01 (Liang: PI) 5/01/2018– 04/30/2023 NIH/NIGMS

MERIT: Models and Algorithms for Beta-Barrel Membrane Proteins and Stochastic Networks

R01CA204962-01A1 (Liang: Lead MPI) 4/01/2017 – 3/31/2021
NIH/NCI
Database and Tools for Functional Inference and Mechanistic Insight into Somatic Cancer Mutations

Citations:

1. Qiu Sun, Alan Perez-Rathke, Daniel M. Czajkowsky, Zhifeng Shao, and **Jie Liang**. High-resolution single-cell 3D-models of chromatin ensembles during Drosophila embryogenesis. *Nature Communications* 205 12(1):205. doi: 10.1038/s41467-020-20490-9 (2021). PMC7794469

2. Alan Perez-Rathke, Qiu Sun, Boshen Wang, Valentina Boeva, Zhifeng Shao and **Jie Liang**. CHROMATIX: computing the functional landscape of many-body chromatin interactions in transcriptionally active loci from deconvolved single-cells. *Genome Biology*. 21:13, 2020. PMC6966897

3. Gamze Gursoy, Yun Xu, Amy Kenter, and **Jie Liang**. Computational construction of 3D chromatin ensembles and prediction of functional interactions of alpha-globin locus from 5C data. Nucleic Acid Research, 2017. 45:11547-11558. doi.org/10.1093/nar/gkx784, PMC5714131

4. Gamze Gursoy, Yun Xu, Amy Kenter, and **Jie Liang**. Spatial confinement is a major determinant of the folding landscape of human chromosomes. *Nucleic Acid Research*, 2014. DOI:10.1093/nar/gku462. PMCID: PMC4117743

B. Positions, Scientific Appointments and Honors

Positions

2018/8-Prsent	Director, Center of Bioinformatics and Computational Biology, UIC
2017/3-10	Visiting Professor, Dept of Mechanical Engineering, University of Tokyo
2014-present	Richard and Loan Hill Professor of Bioengineering, Richard and Loan Hill Dept. of Bioengineering, UIC
2007-present	Professor of Bioengineering, Dept. of Bioengineering, UIC
2006	Visiting Professor, Institute of Systems Biomedicine, Shanghai Jiaotong University
2003-2007	Associate Professor. Depts. of Bioengineering, and adjunct in Computer Science, UIC
1999-2003	Assistant Professor. Depts. of Bioengineering, Electric Engineering and Computer Science University of Illinois at Chicago, Chicago, IL
1997-1999	Investigator, Dept. of Cheminformatics, SmithKline Beecham Pharm., King of Prussia, PA.

Scientific Appointments

2018,2019	Conference Chair/Co-Chair, IEEE Biomedical and Health Informatics Conference
2015-Present	Associate Editor, IEEE Transaction on Biomedical Engineering
2014-Present	Associate Editor, IEEE J of Biomedical and Health Informatics
2012-2014	Area Editor (Structural Bioinformatics), ISMB Conference
2008-Present	Editorial Board Member, J Computer Science and Technology
2004-2011, 2014-2015	Theme Chair and Co-chair, Bioinformatics and Computational Biology. IEEE Engineering in Medicine and Biology Society (EMBS), Annual International Conference, 2004, San Francisco; 2005 Shanghai, 2008, Vancouver; 2009, Minneapolis; 2011, San Diego, 2014, Chicago, and 2015, Milan.
2000-Present	NIH BDMA study section regular member (2006-2009), NIH MSFD (2019-2022) and MSFC study section (2022-present) regular member *ad hoc* member of Special Emphasis Panels) and NSF (MCB, Math Biology, and DBI).
1996 – 1997	Conference Chair/Co-Chair, IEEE Biomedical and Health Informatics Conference NSF Institute of Mathematics and its Applications: Visiting fellowship for the Special Year of Mathematics in High Performance Computing

Honors

2021	UIC Distinguished Professor
2017	Fellow, Japan Society for the Promotion of Science (JSPS)
2013	Faculty of the Year, College of Medicine, University of Illinois at Chicago
2010-2012	University Scholar, University of Illinois at Chicago
2008	Fellow, American Institute of Biological and Medical Engineering
2004, 2006	Faculty Research Award, College of Engineering, University of Illinois at Chicago

2002	NSF CAREER Award
1995	Best Paper Award. Hawaii Intl Conference on System Sciences. Biotechnology Computing
1994-1996	NSF Postdoctoral Research Associate Fellowship (CISE Division)

C. Contributions to Science

We developed the Constrained Self-Avoiding Chromosome (C-SAC) model and the geometric sequential importance sampling technique, which enabled the elucidation of the physical basis of the scaling properties of chromosomes observed in FISH and Hi-C studies. My laboratory has also developed algorithms for reconstruction of ensemble 3D structures of single-cell chromosomes from population Hi-C and for identifying many-body chromatin interactions.

a. Qiu Sun, Alan Perez-Rathke, Daniel M. Czajkowsky, Zhifeng Shao, and **Jie Liang**. High-resolution single-cell 3D-models of chromatin ensembles during Drosophila embryogenesis. *Nature Communications* 205 12(1):205. doi: 10.1038/s41467-020-20490-9 (2021). PMC7794469
b. Alan Perez-Rathke, Qiu Sun, Boshen Wang, Valentina Boeva, Zhifeng Shao and **Jie Liang**. CHROMATIX: computing the functional landscape of many-body chromatin interactions in transcriptionally active loci from deconvolved single-cells. *Genome Biology*. 21:13, 2020. PMC6966897
c. Gamze Gursoy, Yun Xu, **Jie Liang**. Spatial organization of the budding yeast genome in the cell nucleus and identification of specific chromatin interactions from multi-chromosome constrained chromatin model. *PLOS Computational Biology, 2017. 13(7): e1005658. PMC5714131*
d. Gamze Gursoy, Yun Xu, Amy Kenter, and **Jie Liang**. Computational construction of 3D chromatin ensembles and prediction of functional interactions of alpha-globin locus from 5C data. *Nucleic Acid Research*, 2017. 45:11547-11558. doi.org/10.1093/nar/gkx784, PMC5714131

1. Stochasticity plays important role in biology. However, computational studies of stochasticity is challenging. While the Gillespie algorithm has been widely used to solve the discrete Chemical Master Equation (dCME), it is ineffective in sampling rare events and cannot be used to infer the full probability landscape of complex network. We developed the finite buffer ACME algorithm for optimal state enumeration by reducing the state space by O(n!) while providing exact solutions to dCME. Our ACME method currently solves the largest stochastic networks exactly (eg. of 16 dimension with 16 nodes). We also developed a theoretical framework based on quotient matrix and stochastic ordering for bounding errors of state space truncation, so accurate characterization of probability landscape can be achieved: one will know if major probabilistic peaks are missing with given computing resources, and whether exact solution is possible. Significant biological insight on robustness of network and roles of network architecture in determining epigenetic states of cell were revealed from computational results using the ACME method.
    a. Youfang Cao, Xue Lei, Ruy M. Ribeiro, Alan S. Perelson, **Jie Liang**. Probabilistic Control of HIV Latency and Transactivation by the Tat Gene Circuit. *Proc Natl Acad Sci USA*, **2018**, 115:12453-12458
    b. Youfang Cao, Anna Terebus, and **Jie Liang**. Accurate Chemical Master Equation solution using multi-finite buffers *SIAM Multiscale Modeling and Simulation*, **2016.** Accepted. DOI:10.1137/15M1034180
    c. Youfang Cao, Anna Terebus, and **Jie Liang**. State space truncation with quantified errors for accurate solutions to discrete Chemical Master Equation. *Bulletin of Mathematical Biology,* **2016.** 78(4):617-661. PMCID: PMC4896403, DOI: 10.1007/s11538-016-0149-1
    d. Youfang Cao, Hsiao-Mei Lu, and **Jie Liang**. Probability landscape of heritable and robust epigenetic state of lysogeny in phage lambda. *Proc Natl Acad Sci USA* **2010**, 26;107(43):18445-50. PMCID: PMC2972968

2. Understanding the relationship of structure and function of proteins is a problem of central importance in molecular biology. I have pioneered the computational approach of inferring biochemical roles of proteins by matching protein local surfaces based on customized similarity measures derived from evolutionary rates, which are inferred specifically for the relevant protein functional surfaces. The Bayesian continuous time Markov model for evolution of protein binding surfaces allows discrimination of selection pressure due to biochemical functions from pressures due to structural constraints, folding stability, and folding kinetics, a long standing challenging problem in molecular evolution. The CastP protein surface server in my lab is widely used in the structural biology community.
    a. Yan-Yuan Tseng, Joseph Dundas, and **Jie Liang**. Predicting protein function and binding profile via matching of local evolutionary and geometric surface patterns *J. Mol. Biol.* 2009, 387(2) *451-464*. PMCID: PMC2670802

b. Yan-Yuan Tseng and **Jie Liang**. Estimation of amino acid residue substitution rates at local spatial regions and application in protein function inference: A Bayesian Monte Carlo approach *Mol Biol Evo. 2006*, **23**:421-436. PMID: 16251508.
c. Joe Dundas, Zheng Ouyang, Jeffery Tseng, Andrew Binkowski, Yaron Turpaz and **Jie Liang**. CASTp: computer atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues *Nucleic Acids Research, 2006*, **34**:W116-118, doi:10.1093/nar/gkl282. PMCID: PMC1538779.
d. T. Andrew Binkowski, Larisa Adamian and **Jie Liang**. Inferring functional relationship of proteins from local sequence and spatial surface patterns. *J Mol Biol. 2003*, **332**:505-526. PMID: 1294849

3. Membrane proteins constitute a large portion of a typical genome. My students and I have pioneered computational studies of □-barrel membrane proteins (□MPs), an important class of membrane proteins, and have made significant contributions in understanding the physical organizing principles of the structures of □MPs. We have developed the state-of-the-art method for predicting the structure of the transmembrane (TM) section of □mps without requiring knowledge of structural template (Naveed et al, 2012). We also currently have the best results in the challenging task of predicting structures of extracellular loops of □MPs. These progresses were made by developing original model and algorithms from ground up. Overall, research from the Liang lab has made significant contributions in understanding the stability principles of membrane proteins, in predicting membrane protein structures, in predicting protein-protein interactions, in designing novel □MPs and engineering their dissociation behavior, in inferring evolutionary pattern of membrane proteins, and in understanding membrane protein-protein interactions. An example is our recent success in designing a novel outer membrane protein OmpGF that is electrophysiologically fully functional (Lin et al, 2017)
   a. Wei Tian, Meishan Lin, Ke Tang, **Jie Liang**, and Hammad Naveed. High-resolution structure prediction of β-barrel membrane proteins. *Proc Natl Acad Sci USA* **2018**; 113(43):11991-11993.
   b. Meishan Lin, Dennis Gessmann, Hammad Naveed, and **Jie Liang**. Outer Membrane Protein Folding and Topology from a Computational Transfer Free Energy Scale. *J Am Chem Soc,* **2016**; 138(8):2592-601. PMCID: PMC4881743
   c. Hammad Naveed, Yun Xu, Ronald Jackups, Jr., and **Jie Liang** Predicting Three-Dimensional Structures of Transmembrane Domains of b-Barrel Membrane Proteins. *J Am Chem Soc* , **2012** 134 (3), pp 1775-1781, DOI: 10.1021/ja209895m. PMCID: PMC3415959
   d. Hammad Naveed, Ronald Jackups, Jr, and **Jie Liang**. Predicting weakly stable regions, oligomerization state, and protein.protein interfaces in transmembrane domains of outer membrane proteins. *Proc Natl Acad Sci USA* **2009**, 106(31):12735-12740 *doi: 10.1073/pnas.0902169106*. PMCID: PMC2722334

4. My work on protein geometry focused on development of algorithm and software for computing metric properties of protein molecules based on alpha shape theory, as well as for computing topological properties such as voids and cavities in proteins based on early ideas of persistent homology. Further development based on the discrete flow and pocket algorithm lead to methods for computing protein binding surfaces. In addition, I developed a boundary element method for solving Poisson-Boltzmann equation of continuum model of molecular electrostatics. A number of important discoveries have also been made, including the general principles of protein packing.
   a. **Jie Liang**, Herbert Edelsbrunner, Ping Fu, P.V. Sudhakar and Shankar Subramaniam. Analytical shape computing of macromolecules I: molecular area and volume through alpha shape. *Proteins, 1998*, **33**:1-17. PMID: 9741840
   b. **Jie Liang**, Herbert Edelsbrunner, and Clare Woodward. Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design. *Protein Science, 1998*, **7**:1884-1897. PMCID: PMC2144175.
   c. **Jie Liang** and Shankar Subranmaniam. Computation of molecular electrostatics with boundary element methods. *Biophys J. 1997*, **73**:1830-1841. PMCID: PMC1181083.
   d. **Jie Liang** and Ken A. Dill. Are proteins well-packed? *Biophys J. 2001*, **81**:751-766. PMCID: PMC1301551

A complete list of publications is available at: https://scholar.google.com/citations?user=dda1fx0AAAAJ&hl=en or www.ncbi.nlm.nih.gov/sites/myncbi/jie.liang.1/bibliography/40537334/public/?sort=date&direction=descending

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: Li, Liming

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

POSITION TITLE: Associate Professor of Biochemistry

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
Sichuan Agricultural University, Sichuan, China	BS	02/1982	Horticulture
Michigan State University, East Lansing, MI	MS	06/1988	Horticulture
University of California, Davis, CA	PhD	12/1992	Biochemistry
Harvard University, Cambridge, MA	Postdoc	10/1997	Circadian Biology
University of Chicago, Chicago, IL	Postdoc	08/2002	Prion Biology

A. Personal Statement

I am an Associate Professor of Biochemistry and Molecular Genetics, and my research is focused on prion biology – protein conformation-based transmission and epigenetics biology. I have a broad training in biology, with a PhD in biochemistry. My laboratory at the Feinberg School of Medicine, Northwestern University, has made several notable contributions in the prion field: we reported that the heat shock factor (Hsf1) and the Hsp70p nucleotide exchange factor (Sse1p) are important cellular factors required for the *de novo* formation, propagation, and variant establishment of the yeast [PSI^+] prion; we have successfully created several distinct PrP- transgenic *C. elegans* lines that recapitulate a number of functional and neuropathological features of mammalian prion models and thus establishing *C. elegans* as a useful model organism for prion research; we identified a novel prion element in the budding yeast, [SWT^+], whose protein determinant, Swi1, is a chromatin remodeling factor that plays a regulatory role in gene expression; we investigated interactions among three prions, [PSI^+], [PIN^+], and [SWT^+]; and we showed that the first 32 amino acid residues of Swi1p is sufficient for propagating [SWT^+] conformation. Utilizing the yeast [SWT^+] prion system discovered in my laboratory, we have established a robust and straightforward high throughput screening platform that has allowed us to identify several novel anti-prion compounds whose anti-prion action mechanisms are currently being studied. I believe I have the ability, knowledge, and leadership to successfully function as a co-investigator in executing this collaborative research project. I have an in-depth understanding of protein structures and functions and an extensive experience in heterogeneous recombinant protein expression, purification, and follow-up biochemical assays. My knowledge and experience will allow me to make unique contributions to this

proposed CBC Catalyst project. I look forward to working together with Drs. Long and Liang on this important and exciting project.

Ongoing and recently completed projects that I would like to highlight include:

R01 GM126318
Du (PI) , Role: co-investigator
08/01/18 – 07/31/23
Prion-mediated protein aggregation/co-aggregation and cellular consequence

R01 GM110045
Li (PI)
07/01/14 - 06/30/18
High Throughput Screening for Small Molecules Facilitating Prion Study.

MCB-1122135 (NSF)
Li (PI)
11/01/11 – 10/31/16
Investigation of the Yeast Prion [SWI^{+}]

Citations:

1. Du, Z., Park, K., Yu, H., Fan, Q. & **Li, L.** (2008) A newly identified prion linked to the chromatin remodeling factor Swi1 in *Saccharomyces cerevisiae*. Nature Genetics 40(4), 460-465. PMCID: PMC4063924
2. Park, K.W., and **Li, L.** (2011) Prion protein in *Caenorhabditis elegans*: Distinct models of anti-BAX and neuropathology. Prion 5(1), 28-38. PMCID: PMC3038003
3. Du, Z., Zhang, Y. & **Li, L.** (2015) [SWI^{+}] abolishes yeast multicellular growth by triggering conformational changes of multiple regulators required for flocculin gene expression. Cell Reports, 13(12), 2865-2878. PMCID: PMC4704862
4. Du, Z., Valtierra, S., Cardona, L.R., Dunne, S.F., Luan, C.H. & **Li, L.** (2019) Identifying anti-prion chemical compounds using a newly established yeast high-throughput screening system. Cell Chemical Biology, 26(12), 1664-1680. PMCID: PMC6984038

B. Positions, Scientific Appointments, and Honors

Positions and Scientific Employment

2011 – Present Associate professor, Department of Biochemistry and Molecular Genetics (previously Molecular Pharmacology and Biological Chemistry), Northwestern University the Feinberg School of Medicine, Chicago, IL
2002 – 2011 Assistant professor, Department of Molecular Pharmacology and Biological Chemistry,
Northwestern University the Feinberg School of Medicine, Chicago, IL
2016 NIH Support for Conferences and Scientific Meetings (R13) applications Review Panel
2015 NIH ZRG1 MDCN-N (03) "Molecular neurodegeneration and Signaling" Review Meeting
2014 Norman Hackerman Advanced Research Program
2012 – Present Editorial Board, *Prion*

2012 – Present	NIH Peer Review Committee: special emphasis panel ZRG1 BST(F) 50, BST(U)50, and BST(C)50: High Throughput Screening assays for drug and probe discovery, ad hoc reviewer
2008 – 2022	*Mail reviewer of National Science Foundation, Division of Molecular and Cellular Biosciences*

Honors

2018	Liz and Eric Lefkofsky Scholar, Lurie Cancer Center, Northwestern University
2004 – 2008	Ellison New Scholar in Global Infectious Disease.
1996	Marlene DeLuca Memorial Prize, (for outstanding research and presentations by young women scientists) in 9th International Symposium on Bioluminescence and Chemiluminescence, Woods Hole, Massachusetts, USA.
1992	Michael Swackhamer Memorial Research Lectureship, UC Davis, CA.
1991 – 1992	Michael Swackhamer Fellowship Award in Excellence in Research, UC Davis, CA.
1990 – 1991	Jastro-Shields Research Scholarship, UC Davis, CA.

C. Contributions to Science

1. My doctoral thesis was conducted in the laboratory of Clark Lagarias at University of California, Davis. There I established an *in vitro* system to allow assembly of functional phytochrome, a red/far-red light receptor essential in regulating plant growth and development. My first postdoctoral training was at Harvard University Professor J Woodland Hastings' laboratory to study circadian regulation using the nocturnal bioluminescence of *Gonyaulax polyedra* (*G.p.*). I then joined Professor Susan Lindquist laboratory for my second postdoctoral training to work on prion biology in University of Chicago/HHMI, where I demonstrated that the prion domain of Sup35, the protein determinant of the yeast prion [*PSI*$^{+}$], is modular and can be transferred to a completely unrelated mammalian steroid hormone receptor, glucocordicoid receptor (GR), to create an artificial prion. This work was published in Science and has been considered as a piece of strong evidence supporting the "protein-only" prion hypothesis. Selective publications from my trainings:

    a. **Li, L.** & Lagarias, J.C. (1994) Phytochrome assembly in living cells of the yeast *Saccharomyce cerevisiae*. Proceedings of National Academy of Sciences of the United States of America, 91(26), 12535-12539
    b. **Li, L.**, Hong, R. & Hastings, J.W. (1997) *Gonyaulax* luciferase contains 3 functional catalytic domains in a single polypeptide chain. Proceedings of National Academy of Sciences of the United States of America, 94(17)**,** 8954-8958.
    c. **Li, L.,** Liu, L., Hong, R., Robertson, D. & Hastings, J.W. (2001) N-terminal intramolecularly conserved histidines of three domains in *Gonylaux* luciferase are responsible for loss of activity in the alkaline region. Biochemistry, 40(6), 1844-1849.
    d. **Li, L.** & Lindquist, S.L. (2000) Creating a protein-based element of inheritance. Science, 287(5453), 661-664.
2. I have continued prion research in my laboratory in Feinberg School of Medicine, Northwestern University. We identified Swi1 as the fourth yeast prion protein. Swi1 is a subunit of SWI/SNF, an ATP-dependent chromatin-remodeling complex that non-covalently modifies and repositions nucleosomes in chromatin and plays an important role in transcriptional regulation. **We showed that Swi1 is** aggregated in [*SWI*$^{+}$] cells but not in non-prion cells, amyloid is the structural basis of [*SWI*$^{+}$], and [*SWI*$^{+}$] is “infectious”. We recently also showed that [*SWI*$^{+}$] sequesters several transcription up-regulators to result in a complete abolishment of yeast multicellularity.

a. Du, Z., Park, K., Yu, H., Fan, Q. & **Li, L.**, (2008) A newly identified prion linked to the chromatin remodeling factor Swi1 in *Saccharomyces cerevisiae*. Nature Genetics, 40(4), 460-465. PMCID: PMC2633598

b. Du, Z., Zhang, Y. & **Li, L.** (2015) The yeast prion [*SWI*$^{+}$] abolishes yeast multicellular growth by triggering conformational changes of multiple regulators required for flocculin gene expression. Cell Reports, 13(12), 2865-2878. PMCID: PMC4704862

c. Du, Z., Regan, J., Bartom, E., Wu, W.S., Zhang, L., Goncharoff, D.K., & **Li, L.**, (2020) Elucidating the regulatory mechanism of Swi1 prion in global transcription and stress responses. Scientific Reports, 10(1):21838. PMCID: PMC7736884

d. Goncharoff, D.K., Cabral, R., Applebey, S.V., Pagadala, M., Du, Z. & **Li, L.** (2021) Defining key residues of the Swi1 prion domain in prion formation and maintenance. Molecular Cell Biology 41(7), e0004421. PMCID: PMC8224238

3. Our laboratory also aimed to establish the non-pathogenic nematode *C. elegans* as an invertebrate model to elucidate the mechanisms underlying the pathogenic conformational switch and its associated toxicity. We successfully established *C. elegans as a model organism for prion research.*

a. Park, K. & **Li. L** (2008) Cytoplastic expression of mouse prion Protein causes severe cellular toxicity In *Caenorhabditis elegans*. Biochemical and Biophysical Research Communicaions, 372(4), 697-702. PMCID: PMC2587115

b. Park, K.W., and **Li, L.** (2011) Prion protein in *Caenorhabditis elegans*: Distinct models of anti-BAX and neuropathology. Prion 5(1), 28-38. PMCID: PMC3038003

c. Nussbaum-Krammer, C.I., Park, KW, **Li, L.** Melki, R. & Morimoto, R.I. (2013) Spreading of a prion domain from cell-to-cell by vesicular transport in *Caenorhabditis elegans.* PLoS Genetics, 9(3), e1003351. PMCID PMC3610634

d. Du, Z., Valtierra, S., Cardona, L.R., Dunne, S.F., Luan, C.H. & **Li, L.** (2019) Identifying anti-prion chemical compounds using a newly established yeast high-throughput screening system. Cell Chemical Biology, 26(12), 1664-1680. PMCID: PMC6984038

4. We explored the influence of heat-shock-factor and molecular chaperones on yeast prion formation and propagation. We demonstrated that they indeed play an important role in prionization.

a. Park, K., Hah, J., Fan, Q., Thiele, D. & **Li, L**. (2006) Heat shock factor regulating yeast prion de novo appearance and strain Inheritance. Genetics, 173(1), 35-47. PMCID: PMC1461444

b. Fan, Q., Park, K., Du, Z., Morano, K. & **Li, L.** (2007) The role of Sse1 in the de novo formation and variant determination of the [*PSI*$^{+}$] prion. Genetics 177(3), 1583-93. PMCID: PMC2147939

c. Crow, E. T. & **Li, L.** (2011) Newly identified prions in budding yeast, and their possible functions. *Seminars Cell Dev Biol,* 22(5), 452-9. PMCID: PMC3155609

d. **Li, L.** & Kowal, A.S. (2012) Environmental regulation of prions in yeast. PLoS Pathogens, 8(11), e1002973. PMCID: PMC3499569

5. We have analyzed prion *de novo* formation and propagation events and investigated the mutual interactions among different prion elements and how such interactions affect each other's appearance and maintenance.

a. Du, Z. & **Li, L.** (2014) Investigations of yeast prions: [*SWI*$^{+}$], [*PSI*$^{+}$], and [*PIN*$^{+}$]. Genetics, 197(2), 685- 700. PMCID: PMC4063924

b. Du, Z., Goncharoff, D.K., Cheng, X. & **Li, L.** (2017) Analysis of [*SWI*$^{+}$] formation and propagation events. Molecular Microbiology 104(1), 105-124. PMCID: PMC5364053

c. Goncharoff, D.K., Du, Z. & **Li, L.** (2018) A brief overview of the Swi1 prion – [*SWI*$^{+}$]. FEMS Yeast Research, 18(6), doi: 10.1093. PMCID: PMC6001882

d. Du, Z., Cho, B. & **Li, L.** (2022) Identifying endogenous cellular proteins destabilizing the propagation of Swi1 prion upon overproduction. Viruses 14(7), 1366 PMCID: PMC9321512

Complete List of Published Work in My Bibliography:

liming li bibliography (nih.gov)