

Kevin Kaichuang Yang

PhD Candidate
Division of Chemistry and Chemical Engineering
California Institute of Technology
1200 E California Blvd
Pasadena, CA, 91125

614 260 9454
yangkky@gmail.com
<http://yangkky.github.io>

EDUCATION

2018 (projected)	PhD in Chemical Engineering	California Institute of Technology
2011	B.S. in Chemical Engineering	Ohio State University

HONORS, AWARDS, AND FELLOWSHIPS

2017	Caltech Chemistry and Chemical Engineering Teaching Assistantship Award
2015	NIH Biotechnology Leadership Pre-doctoral Training Program
2013	National Science Foundation Research Experience for Teachers Grant
2013	Green Dot Public Schools Rising Star Teacher
2011	National Science Foundation Graduate Research Fellowship

PUBLICATIONS

Peer-Reviewed Papers

3. **Yang, K. K.**, Wu, Z., Bedbrook, C. N., Arnold, F.H. "Protein embeddings for machine learning." *Bioinformatics*. 23 March 2018. doi.org/10.1093/bioinformatics/bty178
2. *Bedbrook, C. N., ***Yang, K. K.**, Rice, A. J., Gradinaru, V., Arnold, F.H. "Machine learning to design integral membrane channelrhodopsins for efficient eukaryotic expression and plasma membrane localization". *PLOS Comp. Bio.* 23 Oct 2017. doi.org/10.1371/journal.pcbi.1005786
1. Bedbrook, C. N., Rice, A. J., **Yang, K. K.**, Ding, X., Chen, S., LeProust, E. M., Gradinaru, V., Arnold, F.H. Structure-guided SCHEMA recombination generates diverse chimeric channelrhodopsins. *PNAS*. 10 Mar 2017. doi/10.1073/pnas.170026911

[*co-first authors]

Conference presentations

1. Oral Presentation: "Machine Learning to Predict Eukaryotic Expression and Plasma Membrane Localization of an Integral Membrane Protein." Proteins Gordon Research Seminar, Holderness, NH. 17 June 2017.

RESEARCH EXPERIENCE

Graduate Research Assistant, Professor Frances Arnold's Group
California Institute of Technology, Pasadena, CA (August 2014 - Present)

- Used Gaussian process models (github.com/yangkky/gpmodel) to design channelrhodopsins with improved properties
- Designed embedded representations of protein sequences based on doc2vec to streamline machine learning pipelines (github.com/fha_lab/embeddings_reproduction)
- Built neural machine translation models on PyTorch to predict signal peptides from their corresponding mature protein sequences

Computational Intern

Ambry Genetics, Aliso Viejo, CA (June 2017 – September 2017)

- Developed and implemented neural network models in Keras and PyTorch to predict outcomes of genetic variation by transferring information across paralogous proteins
- Incorporated model into a pipeline that finds paralogs for variants of interest and then uses paralogs and model to predict variant outcomes

TEACHING AND MENTORING

Teaching Assistantships

2018	Caltech ChE/BE 163 Introduction to biomolecular engineering
2016	Caltech ChE/BE 163 Introduction to biomolecular engineering
2015	Caltech ChE 101 Chemical reaction engineering

Physics and Math Teacher

Green Dot Public Schools, Los Angeles, CA (August 2011 – August 2014)

- Planned and implemented curricula for algebra 1 and physics
- Received highest evaluation based on test scores and observations