Investigating the genotype to phenotype link for a visual protein

Steven K. Chen[1], Jing Liu[1], Belinda S. W. Chang[1,2,3]

Department of Cell & Systems Biology[1],
Department of Ecology & Evolutionary Biology[2], Centre for the Analysis of Genome Evolution & Function[3]
University of Toronto, 25 Harbord Street, Toronto, Ontario, Canada, M5S 3G5

Introduction

As DNA sequencing becomes increasingly routine in clinics for diagnosis and prescribing treatment of diseases, the problem of accurately characterizing the effects of genetic variants is becoming ever more important. Classification standards, such as the American College of Medical and Genomics standard, have been proposed to classify genetic variants ranging from "pathogenic" to "benign"[1]. However, the effects for most variants are unknown and have thus been classified as "variants of unknown significance" (VUS). Even more poorly understood are the gene products that manifest from these variants and the mechanisms by which they carry out their effects.

In the visual system, genetic variation in the *RHO* gene which encodes for the visual pigment rhodopsin is the most common cause of retinitis pigmentosa (RP), a degenerative retinal disease. Over 150 different mutations in *RHO* have been identified[2] and more are being discovered every year. However, many of these variants are VUS because our ability to understand their effects is being outpaced by the rate that new variants are being discovered.

To keep pace with the increasing number of VUS requires a systematic approach that can be easily scaled and is able to classify variants accurately, which is difficult to achieve with a light activated protein. Recent advances in sequencing technologies, however, have accelerated the development of deep scanning mutagenesis, a promising solution to this dilemma that offers the potential for rapid classification of the effects for thousands of variants in a single experiment.

In this study, we aim to develop and apply a deep scanning mutagenesis approach that combines fluorescence-activated cell sorting and deep sequencing to help classify rhodopsin variants and to determine the mechanisms by which these variants confer their effects.

An engineered yeast system for measuring rhodopsin activation

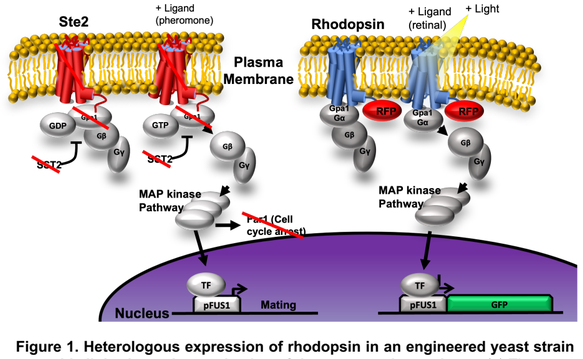

Figure 1. Heterologous expression of rhodopsin in an engineered yeast strain to enable light-dependent activation of the yeast mating pathway. (a) The yeast mating pathway. (b) Rhodopsin activates an engineered yeast mating pathway which activates a GFP fluorescent reporter. Rhodopsin is also fused to ymScarlet fluorescent protein to track its expression and localization in yeast cells.

A variant library for testing all possible missense variants within a functionally important region in rhodopsin

- The variant library comprises of 1,976 missense variants where each of the 104 sites in the target region was mutated to all 19 other amino acids
- Extracellular loop 2 (EL2) is a major hotspot where RP mutations are known to cluster
- Helix 5 (H5), Intracellular Loop 3 (ICL3), and Helix 6 (H6) undergo pronounced conformational shifts during receptor activation
- ICL3 are within contact distance with downstream Gα G protein

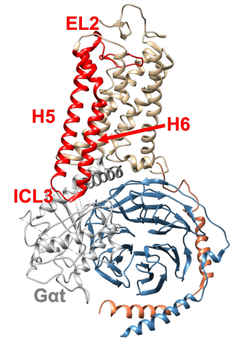

Figure 2. Structure of rhodopsin in complex with G proteins.

Rhodopsin activation and protein abundance varies widely among variants

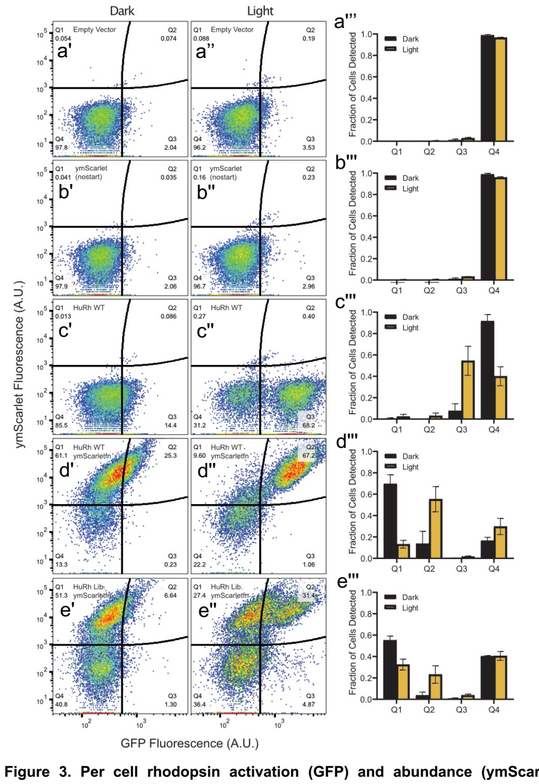

Figure 3. Per cell rhodopsin activation (GFP) and abundance (ymScarlet) under light vs. dark conditions. (a) Engineered yeast not expressing rhodopsin in the dark (a'), light (a''), and fraction of cells detected in each quadrant (a'''). **(b-e)** Same as in (a) except (b) contains ymScarlet with no start codon, (c) is expressing wild type human rhodopsin (HuRh WT), (d) is expressing wild type human rhodopsin fused to ymScarlet (HuRh WT ymScarletfn), and (e) is expressing the library of rhodopsin variants fused to ymScarlet (HuRh Lib. ymScarletfn).

For most variants, sufficient receptor abundance is required to elicit detectable signaling

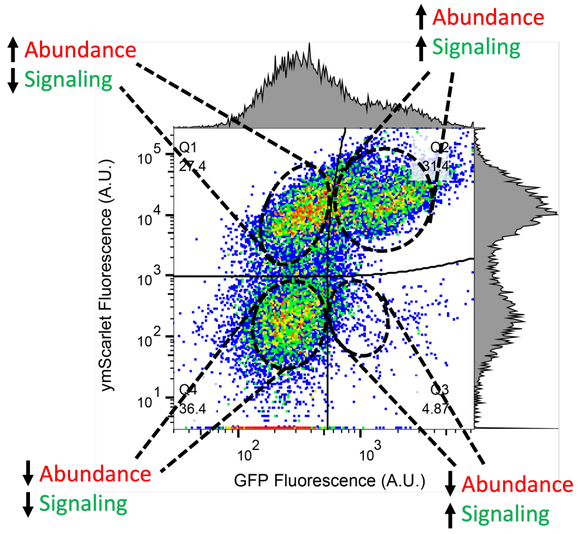

Figure 4. Two-dimensional plot of rhodopsin activation (GFP) and per cell protein abundance (ymScarlet) when our variant library is expressed and light activated in our yeast system.

Imaging rhodopsin in yeast to distinguish misfolded variants from those that fold but do not signal

- Misfolded proteins in yeast are known to form aggregates[3]
- Can we use differences in membrane heterogeneity to distinguish misfolded variants from those that fold but do not signal?

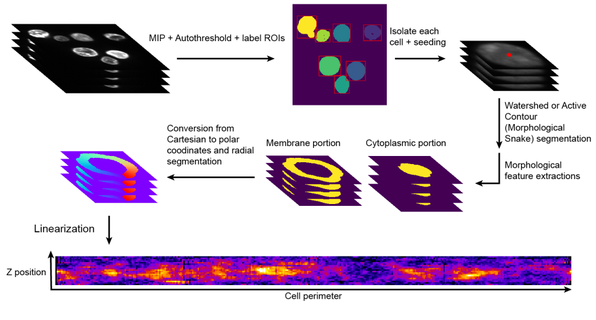

Figure 5. Image processing workflow to segment the plasma membrane of yeast cells that are expressing rhodopsin variants. Maximum intensity projection of the input image is generated, used to create a binary mask of the image by thresholding. Each object in the binary mask is labelled and isolated using its bounding rectangle. Each of the isolated cells is segmented and the resulting segmentation results are separated into binary masks of membrane or the cytoplasmic regions. The membrane mask undergoes further processing where it is linearized and stacked along the depth of the cell, creating a 2D visualization of the protein distribution at the surface.

Well-characterized misfolded variants are heterogeneously expressed on the cell surface

- P23H and V81del amino acid substitutions are known to cause rhodopsin to misfold[4] while N200I is characterized to be superactive (Chen et al. in prep)

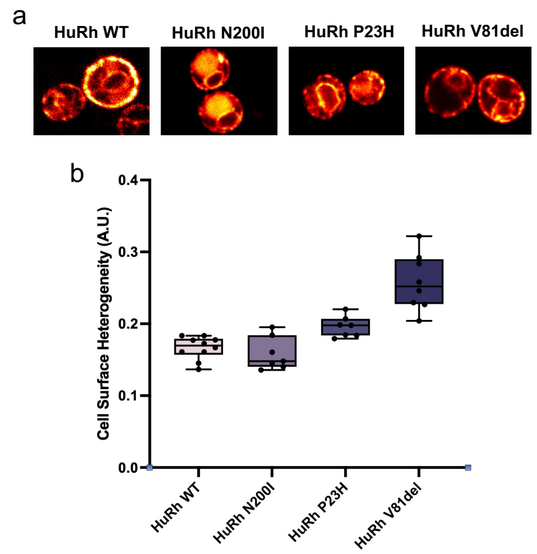

Figure 6. Cell surface heterogeneity for well-characterized variants. (a) Representative confocal micrographs for wild type human rhodopsin (HuRh WT) and three human rhodopsin variants carrying N200I, P23H, and V81del substitutions respectively. (b) Cell surface heterogeneity measurements for the variants in (a).

Future directions

- Use the FACS-Seq approach to complete a deep mutational scan of the library presented here
- Image more variants to test whether differences in membrane heterogeneity can be used to distinguish misfolded variants from those that fold but do not signal

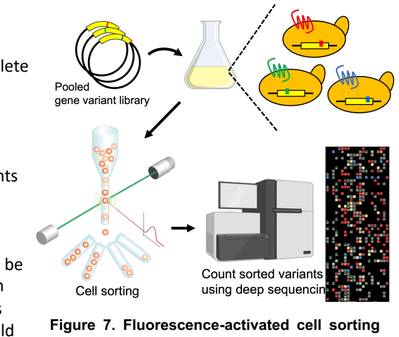

Figure 7. Fluorescence-activated cell sorting coupled to deep sequencing (FACS-Seq) workflow for the deep mutational scan of protein variants.

References

1. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015 May;17(5):405-24.
2. Athanasiou, D., Aguila, M., Bellingham, J., Li, W., McCulley, C., Reeves, P. J., & Cheetham, M. E. (2018). The molecular and cellular basis of rhodopsin retinitis pigmentosa reveals potential strategies for therapy. *Progress in retinal and eye research*, *62*, 1–23. https://doi.org/10.1016/j.preteyeres.2017.10.002
3. Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ, Amon A. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science. 2007 Aug 17;317(5840):916-24.
4. Scott, B. M., Chen, S. K., Bhattacharyya, N., Moalim, A. Y., Plotnikov, S. V., Heon, E., Peisajovich, S. G., & Chang, B. (2019). Coupling of Human Rhodopsin to a Yeast Signaling Pathway Enables Characterization of Mutations Associated with Retinal Disease. *Genetics*, *211*(2), 597–615.

NSERC CRSNG
VISION SCIENCE RESEARCH PROGRAM
UNIVERSITY OF TORONTO FACULTY OF ARTS & SCIENCE