Our proposal **Dual transcriptional programs coordinate lipogenic and membrane stress responsive programs in *C. elegans*** (R01AG068670) explores two major questions related to an understudied membrane stress response, the UPR^{Golgi} and aging. First, we are investigating how the UPR^{Golgi} is linked to aging associated processes such as secretion of yolk proteins in *C. elegans* and links between lipids, SREBP proteins and membrane stress responses in cells from human lipodystrophy patients. Second, we are asking how this branch membrane stress contributes to aging in long-lived *sams-1* animals. For the first goal, we are testing the hypothesis that deficiencies in phospholipid species in the bacterial diet of *C. elegans, E. coli*, sensitize these animals to membrane stress. We obtained engineered *Pseudomonas* strains, which contain a broader complement of phospholipids to mimic E. coli and detected induction of the UPR^{Golgi}. Next, in our studies of the transcriptional regulators of the UPR^{Golgi}, we have identified the specific human orthologs that respond to induction of this stress. Further study of the UPR^{Golgi} in human cells are underway and will be explored in the remainder of the grant. Finally, for our investigation of how responses to membrane stress might change in long-lived *sams-1* animals, we chose to focus on data from an unbiased assay, an RNA seq to examine stress-responsive gene expression in aging *sams-1* animals. This data has provided additional links between genes implicated in the UPR^{Golgi} which we will explore in the remaining time in the grant.

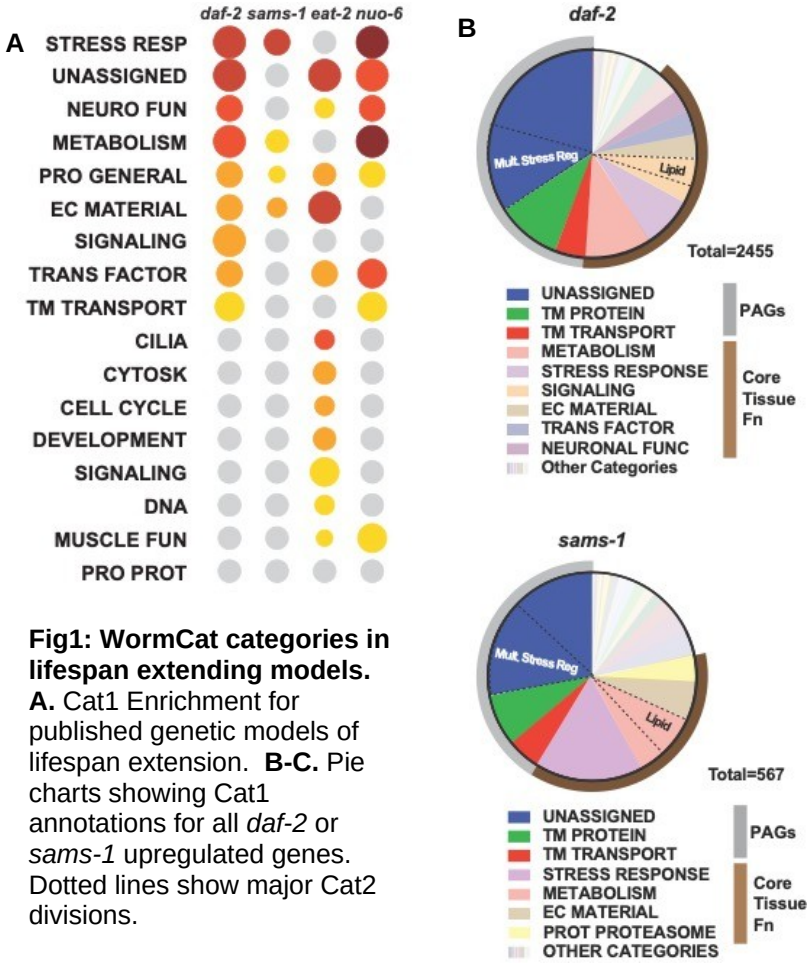

Fig1: WormCat categories in lifespan extending models. **A.** Cat1 Enrichment for published genetic models of lifespan extension. **B-C.** Pie charts showing Cat1 annotations for all *daf-2* or *sams-1* upregulated genes. Dotted lines show major Cat2 divisions.

In our RNA seq studies of aging *sams-1* animals, we used our WormCat annotation and visualization tool to analyze these experiments and saw multiple stress-response pathways increase as these long-lived animals aged. WormCat provides a unique tool for genome-scale analysis, as it includes all *C. elegans* transcripts, allowing inclusion of poorly annotated genes (PAGs) in category analysis. Using WormCat, we noted that *sams-1* animals contained genes from “Unassigned” and “Transmembrane (TM) Transport” categories enriched in both young adult and aging animals and that the TM transport category enrichment increased during aging. Many of these genes are also increased in other genetic and pharmacological models of aging (**Figure 1**). We also made use of a recently published single cell RNA seq data set (Roux, et al (ref)) and asked if PAGs appeared in specific tissue clusters during *C. elegans* aging. Indeed, PAGs from the UNASSIGNED/Unknown category or the UNASSIGNED genes with membrane spanning regions localize to specific clusters associated with neurons, the hypodermis or muscle. This association with well characterized genes within tissue specific groups suggests it is possible to link PAGs to genes with phenotypically tractable assays. While PAGs may be associated with a variety of biological processes, the enrichments UNASSIGNED, TM TRANSPRORT and TM PROTEINS with in multiple genetic (**Fig1**) or pharmacological (Holdorf) models of aging suggests at least some of these proteins could function in lifespan-related processes. **We hypothesize that PAGs regulated during the aging process contribute to the robustness in lifespan regulation and stabilize phenotypes across variable environments.** In order to identify PAGs with roles in aging, we will use artificial intelligence/machine learning tools to cluster PAGs with genes known to function membrane stress pathways, a core-aging related phenotype. Next, we will use quantitative assays to analyze membrane stress phenotypes in aging-sensitized strains with sufficient power to detect effects on robustness.

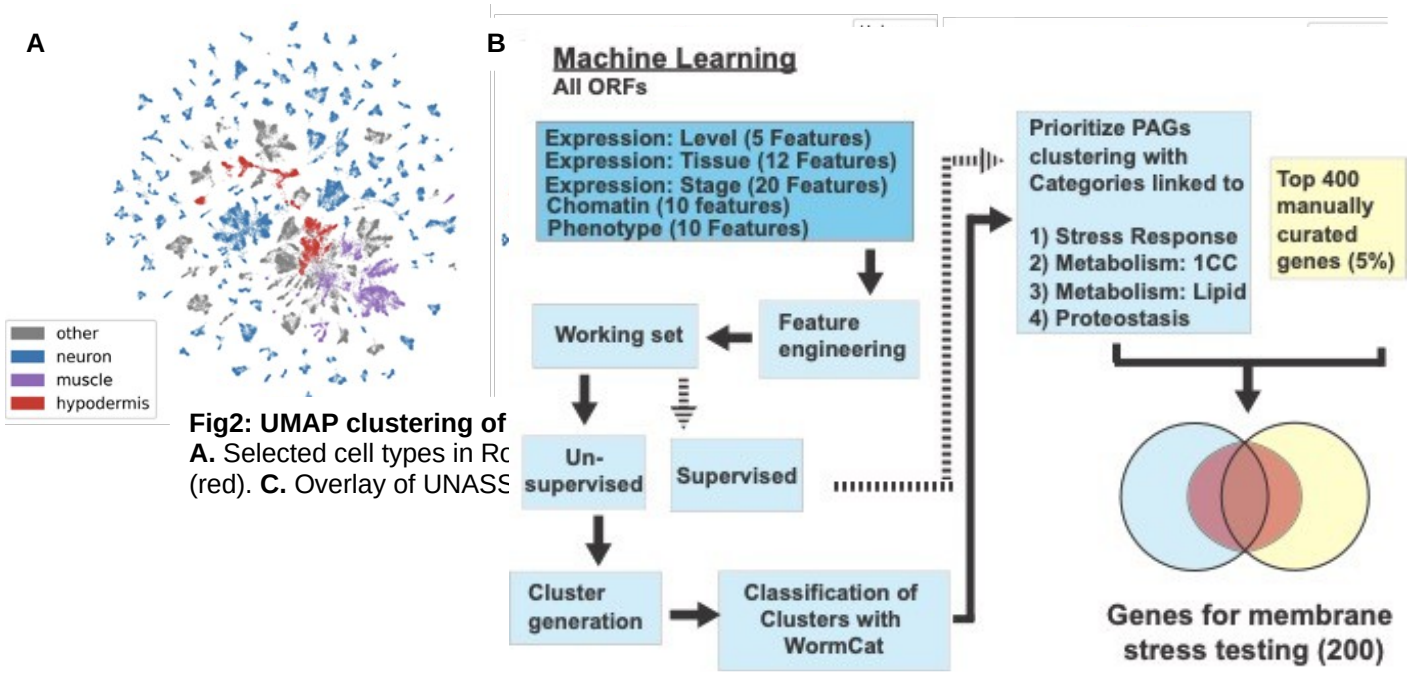

Fig2: UMAP clustering of
A. Selected cell types in Rc
(red). **C.** Overlay of UNASS

Fig3: Flowchart illustrating Machine Learning pipeline. Dashed lines connect alternative approaches.

Two decades of research in the aging field have revealed seven major cellular pathways for regulating aging: IGF/Insulin, mTORC signaling, Sirtuins, AMPK, 1 Carbon Cycle/Epigenetic regulation, Mitochondrial function/hypoxia, and Proteostasis/Autophagy. Many of these pathways connect through sensing and utilization of nutrients or as accessory pathways to buffer the stresses produced by metabolism. While these major pathways have clear links to aging under laboratory conditions, effects on lifespan in natural environments may be variable and complex. Thus, our efforts to discover how these poorly annotated genes function in membrane stress will complement the studies in our R01 and could show how PAGs and also may discover how genes outside the “usual suspects” impact aging or impact core aging related pathways.

Goal: Identify PAGs contributing to membrane stress responses linked to aging

All sequenced organisms, including humans, contain large numbers of genes whose functions have not yet been identified, although they may be conserved across species. Genetic or -omics screens may have failed to reveal phenotypes due to redundant functions with other genes or importance outside the laboratory environment. Another possibility is that some of these genes are important for phenotypic robustness. Robustness is a quality that enhances the stability of a trait in variable conditions (ref). For example, a gene supporting robustness in the aging process might be important for *daf-2* dependent increases in lifespan across variable temperatures or changing nutritional conditions. Therefore, as potential aging-modifiers move out of the laboratory and into wider populations, it is critical to identify genes that change how they act in less controlled conditions. One challenge for identifying which PAGs might act in aging is linking each gene to precise phenotypic assays needed to identify effects on robustness. We hypothesize that artificial intelligence/machine learning tools (AI/ML) can provide a mechanism to group PAGs with genes known to function in lifespan extension. Next, we will examine the effects of PAG knockdown on a panel of membrane stress reporters that change during the aging process. Though utilization of highly quantitative approaches in sensitized background, we will identify those PAGs that alter reporter activation in genetic (*daf-2* and *sams-1*) or pharmacological (rapamycin) models of aging.

Development of Feature lists for Candidate selection and Machine Learning

Genes that are co-expressed may share phenotypes, however, gene expression data contains large amounts of noise, and it may be difficult to identify which expression clusters have the tightest links to functions. Manual construction of co-expression networks is time consuming, error prone and may focus on a limited number of data sources. Machine learning approaches may incorporate large data structures, allowing detection of interactions unlikely to be identified manually (**Fig3**). The first step in both of these approaches is development of a **feature list** from publicly available sources that links each *C. elegans* gene to a matrix of descriptors for expression during time (development), space (tissue) or level along with information on chromatin state (modEncode) and phenotype. Not all of the information in the feature list may be important for the pattern

detection in the machine learning, therefore the next step is **feature engineering** where the data is probed to determine which characteristics are most important to the larger structure (**ref**). There are multiple alternatives to machine learning, each with particular strengths. Algorithms that perform unsupervised learning identify clusters in data without the bias of an internal structure (**ref**). Supervised approaches start with by providing a structure, training a subset of the data with that information, then using that knowledge to probe the larger dataset. Our first approach will be to use unsupervised learning to generate gene clusters based on our feature list. Then we will use WormCat to classify clusters and choose PAGs that are present with clusters of genes linked to aging and stress. An alternative approach would be to use WormCat category information from aging related RNA seq data (**Fig1**) to train a subset of the feature list, then use that to identify similar gene groups likely to be linked to aging from the entire genome.

We will also use multiple techniques to validate this data before phenotype testing. For unsupervised learning, we will use WormCat to analyze a random set of genes of the median cluster size, determining if specific categories are returned. Genes from these lists may also be used as validators in phenotypic assays. We will validate supervised approaches by determining if candidate gene lists are more likely to contain genes with aging-related phenotypes from WormBase.

High-throughput analysis of AI/ML candidates

We will validate AI/ML generated candidates using RNAi to determine effects on a panel of stress reporters in *C. elegans* in wild type and extended lifespan models. Since we expect these candidates may affect robustness, we will perform assays in technical quadruplicates and biological triplicates to assure sufficient power to detect partial effects. Animals will be imaged on the Leica Confocal in the Walker lab and the proprietary software used for quantitation. By comparing ER stress (*hsp-4*/BiP/GRP78 family member::GFP) and Golgi stress markers define in our lab (*warf-1*::GFP), to markers for mitochondrial stress (*hsp-6*::GFP), pathogenic stress (*irg-1*::GFP) and oxidative stress (*sod-3*::GFP), we can distinguish membrane stress responses from other general stresses. RNAi constructs for control for each of these well-described stresses will be included in each assay. We will perform these studies in wild type animals, as well as in *daf-2(e1370), sams-1(lof)* mutants and animals treated with rapamycin. *daf-2* in an ortholog of the insulin receptor with robust effects on lifespan and stress resistance[26]. *sams-1* encoded a S-adenosylmethionine synthase that is essential for phosphatidylcholine production in *C. elegans* [27] and membrane perturbations affect ER[27], Golgi [28] and mitochondrial stress responses [29]. Rapamycin affects the mTORC system[30], extends lifespan across multiple models[31,32] and activates a cohort of PAGs in *C. elegans*[11,33]. Next, we will confirm the GFP reporter assays with more detailed studies, using immunostaining to confirm nuclear accumulation of ER stress effectors such as ATF-6[34] or LET-607/CREB3L3[35] for the Golgi stress response. We will also validate phenotypes in strains with mutant alleles, when available from *C. elegans* community resources, or use CRISPR to create deletion alleles on our lab to confirm RNAi results are specific to the gene of interest.