Data Management Plan

Data and biological materials to be generated by the proposed project will mainly consist of images (e.g. confocal and transmission electron micrographs), results from various types of protein and DNA analyses, mass spectra and analyses of these spectra (proteomics data), DNA constructs, antibodies raised against specific maize proteins, transgenic maize lines expressing various fluorescent fusion proteins, and other seed stocks. New sequences will not be generated, but our results may provide information leading to improvements/corrections in maize gene annotation (gene structure, functional classification of gene product), which will be communicated upon discovery to annotation curators at maizesequence.org and/or maizegdb.org. All data will be stored in digital form, either in the format in which it was originally generated (i.e. Metamorph files, for confocal images; Spectrum Mill files, for mass spectra with results of mass spectra analyses stored in Excel files; tiff files for gel images; Filemaker Pro files for genetics records), or will be converted into a digital form via scanning to create tiff or jpeg files (e.g. western blots or other types of results). All such digital files will be backed up and stored on external drives and/or CDs/DVDs to ensure long-term preservation and accessibility. Data will be summarized in publications and oral presentations, with representative images and other results displayed. At the time of publication, research results will also be summarized at the PI's website http://www.biology.ucsd.edu/labs/smith/ and pdf versions of publications will be made available there if allowed by copyright agreements (and in manuscript form if not).

More specific and detailed information regarding the generation, accessibility and distribution, and long-term storage/maintenance, of biological resources and data to be generated by this project is as follows:

Seeds

Seed stocks to be generated by this project are of two main types: maize mutants in various genetic backgrounds and expressing various fluorescent fusion proteins, and wild type transgenic maize lines expressing fluorescent fusion proteins that we have proposed to generate for Objectives 3A and 4D. All seed stocks will be generated by the end of the three year project period and some earlier (e.g. transgenic lines expressing membrane trafficking markers fused to YFP will be generated during the second year). Bulking and characterization of these lines will be carried out by project personnel with financial support from this grant, with seeds stored during the project period and thereafter at the UCSD Biology Field Station under low temperature and low humidity conditions that preserve seed viability for approximately 20 years. All seed stocks generated during the course of this project will be available for distribution by the end of the 3 year project period and in some cases earlier: all transgenic lines will be made available as soon as they reach the T2 generation (after two backcrosses of the original transformants to B73), and other seed stocks will be made available at the time they are described in publications. In addition to publicizing these resources in publications and talks given by the PI, to further publicize the availability of transgenic lines, images of fluorescent fusion protein localizations will be displayed at the PI's website http://www.biology.ucsd.edu/labs/smith/ and also at the website for David Jackson and Anne Sylvester's NSF Plant Genome Program project where images for a wide variety of other fluorescent fusion protein transgenic maize lines are displayed (http://maize.jcvi.org/cellgenomics/index.shtml). When available as detailed above, seeds will be distributed by the PI upon request without restrictions on re-use, re-distribution, or generation of derivatives except those needed to comply with USDA regulation of transgenic maize (USDA-APHIS permission for shipment to the individual requesting the transgenic seeds). At the conclusion of the PI's research program (e.g. upon retirement or loss of the infrastructure needed to continue storing seeds under suitable conditions), seed stocks generated by this

project of value to the public will be transferred to the Maize Genetics Stock Center in Urbana, Illinois, which will continue to store and distribute seeds upon request for the remainder of their useful life (approximately 20 years from the time of generation), so long as this stock center remains in existence.

Proteomics data

Objective 4 will involve generation and analysis of maize peptide mass spectra for two purposes: comparison of total peptides and phosphopeptides from membrane-associated proteins of wild type vs. *pan* mutants, and identification of proteins co-immunoprecipitating with PAN2. This work will be done in collaboration with our UCSD colleague Steven Briggs, who will generate peptide mass spectra from protein samples provided by the PI. Comparative analysis of membrane-associated peptides and phosphopeptides in wild type vs. pan mutants Is almost complete already and will be finished during the first year of the project. Analysis of proteins co-immunoprecipitating with PAN2 will be completed during the second year of the project. Mass spectra generated in the course of this work will be stored as Spectrum Mill files on external backup drives. At the time of submission of the data for publication (and no later than the end of the three year project period), mass spectra will be deposited at Tranche (https://proteomecommons.org/) for access by the public with the assigned Tranche hash included in the manuscript and reported at the PI's website (http://www.biology.ucsd.edu/labs/smith/). Tranche is an open access repository designed for storage and access of protein mass spectra, which currently has no limits on the quantity of data stored there and does not charge a data storage fee. The stated intention of Tranche is to maintain this resource indefinitely. Spectrum Mill software will be used to search mass spectra against the most recent version of the maize protein database (at maizesequence.org) to identify peptides and their phosphorylation sites, with results stored in Excel files backed up by the PI. These Excel files detailing the Spectrum Mill Interpretation of mass spectra for each peptide (spectral counts, significance scores for peptide identification, iTRAQ label ratios reporting differences in peptide or phosphopeptide abundance in wild type vs. mutant samples, etc.) will be submitted for publication as supplementary data together with explanations of these data and results of related experiments in the manuscript. Upon publication, these Excel files will then be accessible (and maintained on a long-term basis) at the publisher's website. If access to the publication is restricted by subscription, then at least during the restricted period, the Excel files will also be made available at the PI's website (http://www.biology.ucsd.edu/labs/smith/) on an unrestricted basis.

DNA constructs. Several plasmid DNA constructs will be generated during the 3 year project period (e.g. fluorescent fusion protein constructs, constructs for testing yeast two hybrid interactions, constructs for RNAi knockdown of candidate PAN2-interacting proteins). Electronic versions of these constructs will be created with MacVector software and backed up on external hard drives and/or DVDs along with other data generated by this project. Plasmid DNA and glycerol stocks for each construct will be stored frozen for the remainder of the PI's research career. Upon publication of a description of each construct, or at the time that the associated T2 transgenic seed stocks are available in the case of fluorescent fusion protein constructs, DNA constructs and their electronic versions will be distributed by the PI upon request.

Antibodies. Antibodies generated during the 3 year project period (e.g. the antibody raised against a unique region of PAN2 and antibodies raised against candidate PAN2-interacting proteins) will be stored at -80 degrees for the remainder of the PI's research career and upon publication of a description of each antibody, will be made available by the PI upon request.