**Aim 3: X-Net Grabber: A Platform for Cross Species Network building and inference.**

**Rationale:** We propose to build the X-Net Grabber, an intuitive web interface that will give biologists access to all the data, tools, and analysis pipelines required to build gene networks based on experimental and/or inferred data. The end user can build both (i) species-specific networks consisting of multiple edge types (multinetworks, for short) and (ii) cross-species “weighted” networks (“weighted” networks, for short). Users can create these plant networks using the tools developed in Aims 1 and 2 and query them using the interface described below. This will lead the experimental biologists to identify candidate networks of genes, which they can experimentally validate.

**Novelty:** Other web-based tools that allow researchers to query and browse plant gene networks made from data-rich species, such as PlaNet (Mutwil 2011) and ATTED-II (Obayashi 2011) offer large pre-calculated networks, which may get updated periodically. By comparison, X-Net allows biologists to (i) create predicted networks for data-poor species, (ii) create networks based on subsets of experiments, and (iii) to create multinetworks, and/or “weighted” networks using data from multiple species. Because these networks are created “on the fly”, X-Net also gives researchers the ability to not only select which datasets to use, but to select parameters such as orthology method and thresholds for multispecies networks.

**The X-Net Platform**: There are two main network analysis functionalities we propose to create in X-Net: 1) the ability to create a species-specific multinetwork for any given species, and 2) the ability to create a multispecies “weighted” network.

(1) **Species-specific multinetwork**: Network interactions can be divided into two types: those that are determined experimentally and those that are predicted. A species-specific multinework is simply the union of all different types of interactions. The interface for a species-specific interface would allow the researcher to choose: edge types, thresholds (e.g. correlation above 0.6), and sources of data. The species-specific network might come from experimental data, based on Aim 1’s InferNet, or from Interolog. For example, (see Fig. 6) a biologist working on Glycine max, may want to use protein-protein information from Arabidopsis would simply

* + 1. Choose Arabidopsis as the source,
		2. Choose Glycine max as the target,
		3. Choose an orthology definition and threshold
		4. Click on the “Run Interolog” button
		5. Receive a link to the created network.

**(2) Multispecies “weighted” network:** The multispecies “weighted” network will allow researchers to combine networks from any number of species into one multi-species network, where the edges and nodes have confidence values based on “weights” determined by the support from multiple species (in the style of Aim 2) (See Fig. 6). To provide this feature to the community, we will create a “***Network Cart***” in VirtualPlant ([www.virtualplant.org](http://www.virtualplant.org)) (Katari et al 2010) to enable plant biologists to be able to store, manage, and refine the networks they create using X-Net. Because the VirtualPlant user community of biologists finds the existing “Gene Cart” feature both intuitive and powerful, we believe that they will perform some very sophisticated queries with their “Network Carts” as well. This feature occurs often in Systems Biology studies as it enables researchers to refine their network analysis and predictions over iterative rounds of data analysis.

**Anticipated results and Challenges:** The main technical challenge is to deliver this response in reasonable time. The feature that will take the longest is inference of networks using paramaters determined in Aim 1. Inferred networks will depend on orthology cutoffs and the orthology method chosen. In general, this can take hours, even using our 50 core processor parallel cluster. On the other hand, once a network is constructed, it can be stored for later use, provided the data sources have not changed. To this end, we cache each constructed network and destroy the cached results only when the relevant data sources have changed. Further, we offer users to choose parameters that will lead to cached networks. For example, suppose multinetworks for Arabidopsis, Rice, Medicago, and Poplar using Interolog inferences have been constructed, once using the orthology cutoffs of X, and a second time using Y. In his case, both those species and those cutoffs will be offered (perhaps among others). If a later user chooses any subset of the species that includes all inference sources and a cutoff of either X or Y, then the cached networks can be used. In its full functionality, X-Net will enable researchers to create gene networks on the fly- using their own experimental data across a wide range of crop species, and to interpret their data within the context and “knowledge” of other plant genomes.

 [Dennis thinks we need more discussion of the interface and what it does. We can drop the technical challenge cache stuff.]