Supplementary Table 1. List using systematic nomenclature of all the 4562 proteins of *S. cerevisiae* that were successfully tagged and whose purification was attempted.

Supplementary Table 2. List using systematic nomenclature of all 2357 *S. cerevisiae* proteins whose purification was successful. For the purposes of the paper and this table, success was defined as the recovery and high-confidence identification of at least one non-background protein. This allowed us to compensate for the minority of cases in which an interacting protein(s) but not the bait was successfully identified by mass spectrometry.

Supplementary Table 3. List of 4087 proteins that were identified during the course of this project. Of these, 4021 were identified with high confidence (>99%) by MALDI-TOF mass spectrometry and/or LC/MS/MS, and an additional 66 were identified with lower confidence (<99%) by MS but were identified as a partner in a high confidence interaction (i.e. they were identified repeatedly by MS as interacting proteins with lower MS confidence, so that the machine learning algorithms identified them anyway as a high confidence partners). The 66 proteins in this list identified only as interaction partners rather than with at least 99% confidence by mass spectrometry were YIL044C: AGE2, YBR132C: AGP2, YHR129C: ARP1, YPR020W: ATP20, YBR003W: COQ1, YHR055C: CUP1-2, YOR180C: DCI1, YNL001W: DOM34, YNL133C: FYV6, YGL254W: FZF1,YDR302: GPI11, YIR038C: GTT1, YOR278:W HEM4, YFL011W: HXT10, YNL078W: JIP1, YBR015C: MNN2, YOL042:W NGL1, YOR104W: PIN2, YMR006C: PLB2, YFR033C: QCR6, YER173W: RAD24,YDR014W: RAD61, YLR453C: RIF2, YHL027W: RIM101, YMR305C: SCW10, YLL011W: SOF1, YGL169W: SUA5, YDR007W: TRP1, YDL064W: UBC9, YPL252C: YAH1, YBR042C, YBR053C, YBR111C, YCR023C, YDL085C-A, YDR061W, YDR063W, YDR262W , YDR320C-A, YDR458C, YER048W-A, YER134C, YFR018C, YGR201C, YGR205W, YHR046C, YHR198C, YIL055C, YIL157C, YJL062W-A, YJR088C, YKL005C, YKR074W, YLR285W, YLR363W-A, YLR376C, YMR290W-A, YNR034W-A, YNR063W, YOL032W, YOR285, YPR148C, YPR172W, YBL071W, YFR024C, and YLR438C.

Supplementary Table 4. List of *S. cerevisiae* proteins, including a number of chaperones, that were identified by MALDI-TOF MS or LC/MS/MS in more than 3% (71) of all the successful protein purifications. For the purposes of identifying protein-protein interactions these proteins were considered non-specific contaminants and were removed from consideration. The listed cytoplasmic ribosomal proteins were also removed. Because of this, interactions of ribosomal proteins with each other in the ribosome or with many translation factors are not present in our protein interaction network. As well, some chaperones have interactions with many proteins in the normal course of their functioning but were removed from our lists of protein-protein interactions.

Supplementary Table 5. List of 2357 protein-protein interactions involving 1210 proteins in the "intersection dataset," along with the confidence scores for protein identification by mass spectrometry. An interaction is listed in this Table if the interacting protein partner was identified in a single purification by both MALDI-TOF mass spectrometry and LC/MS/MS, and even if the confidence for protein identification by mass spectrometry was relatively low, as long as it was greater than 70%. As described in Supplementary Information, the Z-scores for protein identification by mass spectrometry were rescaled to a scale from 0 to 1 such that a score of 0.5 represents 70% confidence.

Supplementary Table 6. List of 5496 protein-protein interactions involving 2186 proteins in the "merged dataset," along with the confidence scores for protein identification by mass spectrometry. This list contains all 2357 interactions in the "intersection dataset," as well as interactions identified repeatedly by only a single method of mass spectrometry (with a confidence threshold of 70%) and interactions identified by a single mass spectrometry method (with a confidence threshold of 70%) when each of the interacting protein partners was purified. As described in Supplementary Information, the Z-scores for protein identification by mass spectrometry were rescaled to a scale from 0 to 1 such that a score of 0.5 represents 70% confidence.

Supplementary Table 7. List of 7123 protein-protein interactions in the core dataset involving 2708 *S. cerevisiae* proteins. These were the highest confidence interactions (minimum 0.273, mean 0.68, median 0.69) identified by the machine learning procedure described in the main text and Figure 2A. Each interaction is listed along with its confidence score. It is these protein-protein interactions that were used to compute the protein interaction network and identify protein complexes as described in the text and shown in Figures 3D and S3.

Supplementary Table 8. List of 14317 protein-protein interactions in the extended dataset involving 3672 *S. cerevisiae* proteins. Each interaction is listed together with its confidence score as determined by the machine learning procedure. This list includes all the protein-protein interactions listed in Table S7 but now the minimum confidence score is 0.101 rather than 0.273. This Table should have approximately 1400 correct interactions among the 7194 interactions that are not listed in Table S7.

Supplementary Table 9. Complete list of all the putative *S. cerevisiae* protein-protein interactions identified in this study, together with their confidence scores as determined by the machine learning algorithms. The vast majority of the interactions listed here that are not listed in Table S8 are incorrect. They have very low confidence scores (usually much less than 0.1) either because the interactions were not identified reproducibly or because the protein identifications by mass spectrometry had low confidence, or both. This list is provided because some of the additional interactions listed here and not in Tables S7 and S8, perhaps as many as several hundred, could be correct.

Supplementary Table 10. List of protein complexes and their component subunits as identified by the Markov Cluster Algorithm. Each of the protein clusters is given a number from 0 to 546. The MCL algorithm does not do well at separating protein complexes that share subunits. At the moment, achieving this aim for a given protein complex is likely to require manual inspection of the data.