Gygi *et al.* reply: We are pleased that our large-scale method for measuring phosphorylation stoichiometries[1] has inspired additional analysis[2]. We had reported stoichiometries for ~5,000 yeast protein phosphorylation sites, only a small fraction of which were highly conserved across 25 fungal species. We noted as a minor point that lowest-occupancy sites appeared more highly conserved than the highest-stoichiometry sites. Tan & Bader[2] propose an explanation for this second observation: the lowest-stoichiometry phosphorylation sites disproportionately occur on high-abundance proteins, which tend to be more highly conserved. Once background protein conservation rates are taken into account, high-stoichiometry sites may be somewhat more conserved than lower-occupancy ones[2].

Although lowest-stoichiometry phosphorylation sites may be over-represented among the most abundant proteins, the complete set of ~5,000 sites[1] is evenly distributed across proteins of all abundances. Furthermore, when we mapped 12,000 phosphorylation sites from an independent quantitative proteomics study[3] onto the yeast proteome, the median phosphoprotein abundance was lower than the median abundance for nonphosphorylated proteins (Fig. 1a). When we divided the yeast proteome into ten groups by abundance, phosphoproteins (not subjected to enrichment methods) were most represented in the lowest-abundance bin (Fig. 1b). Similarly, phosphoproteins in the lowest-abundance bin generally had more phosphorylation sites (Fig. 1c); notably, the most heavily phosphorylated protein in this dataset, SEC16, contained 87 phosphorylation sites and was in the lowest-abundance bin.

We generally agree with Tan & Bader[2] that their analysis suggests that highest-stoichiometry phosphorylation sites are more conserved than lower-occupancy sites after correction for background conservation rates: although phosphorylation sites in the three lowest stoichiometry classes cluster together, making it difficult to distinguish trends among them, the highest-stoichiometry class was shifted relative to the others, suggesting increased conservation. Although this increased conservation of high-stoichiometry phosphorylation sites is noteworthy, differences in divergence rates were small; thus, in our view the most important observation is that the vast majority of both high- and low-stoichiometry sites are not conserved. This is illustrated in Figure 5 of the original publication[1], and Tan & Bader[2] have confirmed this: according to their analysis, even high-stoichiometry sites diverge at 80–90% of the rate for unmodified residues. Similarly, divergence rates for mid- and low-stoichiometry groups are frequently above 1.0, suggesting by their analysis that these sites are more likely to diverge than structurally similar, yet unmodified, serines, threonines and tyrosines[2].

These studies raise some fascinating issues. Phosphorylation networks are universal and highly conserved features of eukaryotic systems; even minor disruptions in phospho-signaling can have deleterious consequences. Yet despite widespread conservation of specific signaling pathways, the majority of thousands of observed phosphorylation sites have not been conserved at rates substantially different from comparable unmodified serines, threonines and tyrosines. This paradox highlights the incredible plasticity of eukaryotic phospho-signaling networks and raises the question of the evolutionary value of phosphorylation, irrespective of phosphorylation-site stoichiometry. We hope these studies will continue to inspire scientific dialog to better understand the evolution of phosphorylation.

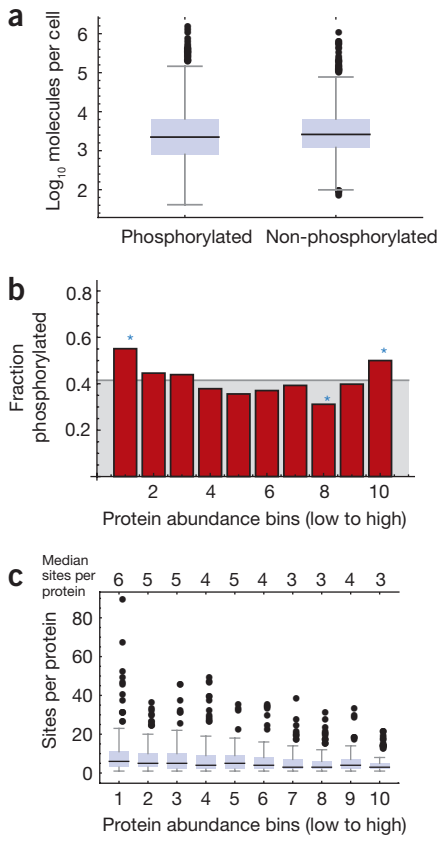

Figure 1 | Phosphorylation versus protein abundance. (**a**) Distributions of 1,413 yeast phosphoproteins and 1,992 nonphosphoproteins[3] with respect to protein abundance[4]. (**b**) Frequencies of protein phosphorylation after dividing all proteins into ten equal-sized abundance bins. Gray shading indicates the overall fraction of proteins that were phosphorylated (0.415). Enrichment in each bin was evaluated using a binomial test with Bonferroni correction for multiple hypothesis testing (*corrected $P < 0.01$). (**c**) Distributions of phosphorylation site counts per protein for proteins in ten equal-sized bins of proteins, classified by abundance[4]. In **a** and **c**, black central lines reflect the median, and boundaries of blue boxes represent 25th and 75th percentile. Whiskers depict the closer of either the most extreme observation in the dataset or the space centered on the median and bounded by a distance of 1.5 times the interquartile range. Dots represent data points falling outside 1.5 times the interquartile range.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Steven P Gygi, Edward L Huttlin, Ronghu Wu, Wilhelm Haas, Noah Dephoure, Mathew Sowa & Bo Zhai

Department of Cell Biology, Harvard Medical School, Boston, Massachusetts, USA.
e-mail: steven_gygi@hms.harvard.edu

1. Wu, R. *et al. Nat. Methods* **8**, 677–683 (2011).
2. Tan, C.S.H. & Bader, G.D. *Nat. Methods* **9**, 317 (2012).
3. Wu, R. *et al. Mol. Cell. Proteomics* **10**, M111.009654 (2011).
4. Ghaemmaghami, S. *et al. Nature* **425**, 737–741 (2003).