charged) phosphatidylserine lipids in the plasma membrane, or do the much less abundant phosphoinositides PIP$_2$ and PIP$_3$ play a key role in targeting?

Heo et al. used the same conjugated form of a constitutively active yeast lipid phosphatase to study the potential role of PIP$_2$ in GTPase–plasma membrane interactions. Using the phosphatase, rather than PLC, to deplete the plasma membrane of PIP$_2$ had little effect on the membrane association of the GTPases (PLC-mediated hydrolysis of PIP$_2$ leads to activation of protein kinase C and Ca$_{2+}$-calmodulin, both of which cause K-Ras to move off the membrane). This was not surprising, because previous work showed that K-Ras binds well to phospholipid vesicles with putative physiological levels (20%) of monovalent phosphatidylserine (4). The surprise came when cells were also treated with inhibitors of phosphatidylinositol 3-kinase to reduce both PIP$_3$ and PIP$_2$ in the membrane: K-Ras4B and the other GTPases with a cluster of basic residues translocated from the plasma membrane. The authors conclude that both phosphoinositides molecules target and anchor clusters of basic residues to the plasma membrane.

This conclusion may generate controversy for two reasons. First, it challenges the common wisdom that PIP$_2$ “is present in negligible amounts in resting cells” (1). Second, it challenges the hypothesis, first put forth by Silvius, that the cluster of basic residues on a GTPase associates with the plasma membrane because its inner leaflet may contain a higher mole fraction of phosphatidylserine than the cytoplasmic leaflets of intracellular organelle membranes, and thus has a more negative surface potential (5). However, data in Yeung et al. (4) support the conclusion that it is PIP$_2$ and PIP$_3$ (3), rather than phosphatidylserine, that likely target K-Ras to the plasma membrane.

Experiments on model phospholipid membranes and theoretical calculations explain why clusters of basic residues in proteins require neither structure nor specific sequences to laterally sequester polyvalent PIP$_2$ and PIP$_3$. These clusters produce a local positive electrostatic potential that extends about 1 nm from the region and acts as a deep “basin of attraction” for multivalent negatively charged phosphoinositides (6). Uncertainty about several factors—for example, the free concentration and lateral distribution of lipids in biological membranes or the proximity of acidic residues to the basic cluster—means that experiments on model membranes could not be used to tease out the relative importance of phosphatidylserine or phosphoinositides in targeting or anchoring peripheral proteins with clusters of basic residues to the plasma membrane. The results on GTPases suggest that the phosphoinositides PIP$_2$ and PIP$_3$ may be more important (3).

The new lipid phosphatase (and kinase) tools should also prove useful for investigating the binding of other proteins with membrane-sticky clusters of basic residues [for example, the scaffolding protein gravid and the protein kinase Src] and the many other plasma membrane processes involving PIP$_2$ and PIP$_3$ (7). Indeed, a different group recently used essentially the same phosphatase tool to investigate how PIP$_3$ affects both endocytosis and the transient receptor potential melanin cluster 8 (TRPM8) Ca$^{2+}$-conducting channel (7). If you are interested in the many functions of the phosphoinositides in the plasma membrane, and how the new tools can be used to investigate these functions, read these important reports (2, 3, 7).

References

Published online 9 November 2006; 10.1126/science.1136314
Include this information when citing this paper.

GENETICS

Delivering New Disease Genes

Lon R. Cardon

Complex diseases represent an extraordinary challenge to geneticists, but recent results are revealing some successful strategies.

Since the human genome was sequenced, advances in human genetics research have steadily built momentum toward identifying genes that influence common human diseases. Validation of millions of genetic variants, rapid advances in genotyping technologies, and the ongoing establishment of repositories of large, population-based patient samples have created expectations of imminent discoveries of prized disease genes. Have these activities truly laid the foundation for gene identification? On page 1461 of this issue, Duerr et al. (1) demonstrate an association between variants in the IL23R gene and Crohn’s disease, a common inflammatory condition of the gastrointestinal tract. Their results show that complex disease genes are finally yielding their secrets and provide crucial validation of the long journey to gene discovery.

Duerr et al. used a genome-wide association approach premised on a simple idea: Assay genome DNA from a sample of cases (diseased patients) and controls for a very large number of genetic variants, or single-nucleotide polymorphisms (SNPs); then, at each SNP site, compare the frequencies among cases and controls (Duerr et al. examined more than 300,000 SNPs). Sites that differ significantly between cases and controls are then validated in independent samples. In practice, of course, genome-wide association is more complicated. Individual differences in most common diseases are thought to arise from a relatively small number of genes (numbering in the tens to hundreds), most of which contribute only modestly to the overall disease risk (2). Therefore, in a large-scale association screen, most disease gene variants are expected to produce only a small “signal” that is difficult to detect among a large number of SNPs. One big question is how many (if any) of the signals will be large enough to separate from the “noise” (3, 4). The answer from the IL23R study is not many, but even a few may be enough to help uncover significant novel disease-associated variants.

The genome-wide association screen of Duerr et al. revealed three SNPs with evidence for disease association more than 10 times as large as that of the next most statistically significant SNP. In genome-wide association scans, such SNPs are usually either artifacts due to genotyping error, or rarely observed examples of variants with large
phenotypic effects. Distinguishing these two outcomes is challenging because there is little statistical and experimental information to guide the process (5). Duerr et al. found that two of the three SNPs with the strongest evidence for association are in a well-known susceptibility gene (CARD15) for Crohn’s disease (6, 7), and the third is a non-synonymous coding change in IL23R, a gene encoding a receptor for the pro-inflammatory cytokine interleukin-23. Thus, two of the top three hits validated the genome-wide association proof-of-principle, and the third was either an artifact or a true positive result.

Replication of genome-wide association findings is essential to distinguish these possibilities. Unfortunately, replication has not been an area of strength in genetic association studies (4). In this case, however, Duerr et al. replicated the IL23R finding unambiguously, revealing strong statistical support for the same SNP, with the same risk allele, and in the same specific phenotypes in larger independent samples of disease patient cases, controls, and families. Together with the recent identification of genes associated with age-related macular degeneration by similar techniques (8, 9), the IL23R finding should help future studies, as researchers can better understand the statistical profiles of genuinely associated disease alleles.

The finding of an association between IL23R and Crohn’s disease has already led to other discoveries. Duerr et al. examined the chromosomal locations of their remaining genome-wide association SNPs to see if others were in the same region as the IL23R gene. Many were, some of which had supportive but not striking statistical evidence for association. Replication studies and further statistical analysis suggest that several of these SNPs contribute independently to Crohn’s disease. These loci of smaller effect were detectable because of the initial discovery of the highly significant variant (see the figure).

Should we be surprised that the genome-wide association revealed only one unambiguous novel SNP for the disease of interest? Probably not. The Crohn’s disease sample size of 547 cases and 548 controls is small for such a study, having the statistical power to detect only large genetic effects, of which there are likely few in the genome. The problem of statistical power is especially salient when the associated variant occurs at low frequency, such as the IL23R variant studied by Duerr et al., which has a frequency of 2 to 3%. Larger sample sizes will be needed to identify more disease loci.

There are at least two important lessons in the strategy used to uncover IL23R and the macular degeneration genes. First, though the initial genome-wide association discoveries may uncover relatively few “low-hanging fruit”—especially with small sample sizes—they may lead to identifying more stubborn variants with smaller effect size (10–12). Second, every SNP counts.

The cumulative effect of all of these aspects of the IL23R study is to lend confidence in the results, which is not always readily apparent in other designs. A recent study of association between the KIBRA gene and human memory (13) adopted an entirely different design, using pooled DNA samples in multistage genotyping, analyzing different memory measures in the primary and replication phases, and studying patient samples from different geographic locations, some with high levels of population substrate. These phenotypic and sampling differences may eventually support the generality of the reported finding, but they will complicate interpretations of external confirmation because the hypotheses generated may be difficult to falsify—negative results could reflect either measurement differences or lack of replication, whereas strictly positive associations would require firm concordance in phenotypes and sampling to demonstrate consistency.

Not all genome-wide association studies will be as successful as the IL23R finding. Sample size will be a key determinant of outcome, as will genetic, population, and phenotypic heterogeneity. In addition, it is increasingly important to present data and results for all analyses conducted, which is one of the few shortcomings of the Duerr et al. report. In any case, results from the current generation of genetic studies should help provide a foundation for the next set of problems, involving detection of rare genetic variants, leveraging the genetics to better understand environmental risk factors, and ultimately, using this hard-won information to improve public health.

References
1. R. H. Duerr et al., Science 314, 1461 (2006); published online 26 October 2006 (10.1126/science.1135245).