(Sma I–Xba I fragment), the *neo* gene (Bgl II–Bam HI fragment), and the *tk* gene (Pvu II fragment) (Fig. 1). The DNA was subjected to electrophoresis in a 0.6% agarose gel, transferred to nylon filters, and hybridized as described [M. Sasaki *et al.*, *Cancer Res.* **49**, 4402 (1989)]. Because the Sma I–Xba I fragment of the Ki-*ras* gene contains repetitive sequences, the probe DNA was prehybridized with a large excess of human placental DNA to block hybridization to the repetitive region of the probe [P. G. Sealey, P. A. Whittaker, E. M. Southern, *Nucleic Acids Res.* **13**, 1905 (1985)].

26. We performed restriction enzyme mapping on human genomic Ki-*ras* DNA that was isolated from a genomic library (LI018; JCRB) by screening with human Ki-*ras* genomic DNA (2.2-kb Xba I–Eco RI fragment).
27. S. Shirasawa and T. Sasazuki, unpublished results.
28. First-strand cDNA was generated from total RNA (10 μg) with oligo dT (21-mer). The Ki-*ras* cDNA was amplified by PCR for 30 cycles (96°C for 60 s, 56°C for 60 s, and 72°C for 40 s) with primers PKF2 and PKR2 (*4*). Primer PKF2 (5′-ATGACTGAATATAAACTTTGTGG-3′) was located in exon 1, starting at initiator codon 1 (ATG) 11 bp downstream of the Stu I site. Primer PKR2 (5′-ACAAAGAAAGCCCTCCCCAG-3′) was located in exon 2, which is 12 kb from exon 1 in the Ki-*ras* gene (*19, 33*). The PCR products were shown to be the predicted length (239 bp) by gel electrophoresis and ethidium bromide staining. The SSO probes for normal (Gly) and mutated (Cys) sequences at codon 12 of Ki-*ras* were Gly^{12}-SSO (5′-GTTGGAGCTGGTGGCGTA-3′) and Cys^{12}-SSO (5′-GTTGGAGCTTGTGGCGTA-3′), respectively. The sequences of the Gly^{13}-SSO and Asp^{13}-SSO are shown in (*24*). Dot blot hybridization of the Gly^{13}-SSO and Asp^{13}-SSO probes to the amplified cDNA fragments was performed as described in (*24*). Hybridization with the Gly^{12}-SSO and Cys^{12}-SSO probes was performed as described in (*4*). Hybridization and washing temperatures were 56° and 59°C, respectively.
29. J. K. V. Willson, G. N. Bittner, T. D. Oberley, F. Meisner, J. L. Weese, *Cancer Res.* **47**, 2704 (1987).
30. Proliferation in semisolid agar was estimated as described in [H. Paterson *et al.*, *Cell* **51**, 803 (1987)] by plating 3×10^3 cells in DMEM containing 10% FCS and 0.33% Noble agar (Difco, Detroit, MI) on top of a 0.66% hard agar layer. Colonies were scored after 3 weeks.
31. For tumorigenicity assays, 10^7 cells suspended in DMEM that contained 10% FCS (0.2 ml) were injected subcutaneously into the flanks of 5-week-old nude mice.
32. Expression was determined by Northern (RNA) blot analysis. Total cellular RNA (15 μg) was separated by electrophoresis in a 1% agarose-formaldehyde gel, transferred to nitrocellulose, and hybridized with a probe for c-*myc* (Cla I–Eco RI fragment in exon 3) as described in (*9*). The nitrocellulose filter was subsequently washed to remove the c-*myc* probe and was rehybridized with a probe for human β tubulin (β2) cDNA [S. A. Lewis, M. E. Gilmartin, J. L. Hall, N. J. Cowan, *J. Mol. Biol.* **182**, 11 (1985)] as an internal control.
33. D. J. Capon *et al.*, *Nature* **304**, 507 (1983).
34. We thank K. Shimizu, S. Aizawa, and J. K. V. Willson for providing Ki-*ras* genomic clones, pMC1DT-A plasmid, and VACO 235, respectively, and M. Katsuki, A. Kimura, and M. Ohara for critical comments. Supported by a Grant-in-Aid for Cancer Research (03152096) from the Ministry of Education, Science, and Culture, Japan.

14 September 1992; accepted 7 December 1992

Phosphatidylinositol 3-Kinase Encoded by Yeast *VPS34* Gene Essential for Protein Sorting

Peter V. Schu, Kaoru Takegawa, Michael J. Fry, Jeffrey H. Stack, Michael D. Waterfield, Scott D. Emr*

The *VPS*34 gene product (Vps34p) is required for protein sorting to the lysosome-like vacuole of the yeast *Saccharomyces cerevisiae*. Vps34p shares significant sequence similarity with the catalytic subunit of bovine phosphatidylinositol (PI) 3-kinase [the 110-kilodalton (p110) subunit of PI 3-kinase], which is known to interact with activated cell surface receptor tyrosine kinases. Yeast strains deleted for the *VPS*34 gene or carrying *vps*34 point mutations lacked detectable PI 3-kinase activity and exhibited severe defects in vacuolar protein sorting. Overexpression of Vps34p resulted in an increase in PI 3-kinase activity, and this activity was specifically precipitated with antisera to Vps34p. *VPS*34 encodes a yeast PI 3-kinase, and this enzyme appears to regulate intracellular protein trafficking decisions.

A variety of peptide growth factors and hormones, such as platelet-derived growth factor (PDGF), nerve growth factor (NGF), and insulin mediate their cellular effects by interaction with cell surface receptor tyrosine kinases. The interaction of these ligands with their cognate receptors induces a series of intracellular signaling events, including stimulation of protein tyrosine kinase activity and the formation and activation of multiprotein complexes at the plasma membrane (*1–3*). These receptor complexes contain enzymes that take part in phosphoinositide metabolism, including phospholipase C (PLC) and PI 3-kinase (*1*). PLC-γ is known to cleave PI 4,5-bisphosphate [$PI(4,5)P_2$] to generate the second messengers diacylglycerol, which activates protein kinase C, and inositol 1,4,5-trisphosphate (IP_3), which can mobilize intracellular Ca (*4*). The association of PI 3-kinase with receptor tyrosine kinases has permitted the identification of a family of phosphoinositides phosphorylated at the 3 position of the inositol ring that appear to act in a signaling pathway distinct from that resulting from turnover of $PI(4,5)P_2$ (*5*). None of the three reaction products of PI 3-kinase—PI 3-monophosphate (PI3P), PI 3,4-bisphosphate [$PI(3,4)P_2$], or PI 3,4,5-trisphosphate [$PI(3,4,5)P_3$]—is a substrate of the known PLC isozymes, suggesting that the lipid compounds themselves function as second messengers in cell signaling (*6, 7*).

Mammalian PI 3-kinase exists as a heterodimer of an 85-kD (p85) and a 110-kD (p110) subunit (*8–10*). The cDNA encoding p110 from bovine brain has been cloned, and the p110 protein has been characterized as the catalytically active subunit of the PI 3-kinase complex (*11*). The p85 subunit is believed to mediate the binding of the p110 subunit to active receptor protein tyrosine kinases through its Src homology 2 (SH2) domains (*12*). This association of p85-p110 with the receptor might lead to the activation of the lipid kinase (*13, 14*). The binding of tyrosine-phosphorylated peptides to the SH2 domains of p85-p110 in vitro stimulates the PI 3-kinase activity fivefold (*2*).

The amino acid sequence deduced from the cDNA sequence of the catalytic subunit of PI 3-kinase from bovine brain is similar to the sequence of the yeast *VPS*34 gene product (*11*), which is essential for the efficient sorting of vacuolar hydrolases in yeast (Fig. 1) (*15*). In eukaryotic cells, vacuolar (lysosomal) enzymes and proteins destined for secretion are transported together through the early stages of the secretory pathway (endoplasmic reticulum and Golgi apparatus) (*16, 17*). Protein transport is mediated by vesicles that bud from one compartment and then dock and fuse with the next compartment in the pathway (*18*). Vacuolar (lysosomal) hydrolases are actively sorted away from proteins destined for the cell surface in a late Golgi compartment (*19, 20*). Mutations in the *VPS*34 gene lead to the missorting and secretion of Golgi apparatus–modified precursor forms of several vacuolar hydrolases, including carboxypeptidase Y (CPY) and proteinase A (PrA) (*15, 17*). Subcellular fractionation experiments have shown that the product of the *VPS*34 gene, Vps34p, is associated through protein-protein interactions with an organelle distinct from the vacuole (*15, 21*).

PI 3-kinase activity is readily detected in yeast cell extracts (*22*). To determine whether *VPS*34 encodes PI 3-kinase, we assayed for PI 3-kinase activity in extracts from wild-type yeast cells, cells with the

P. V. Schu, K. Takegawa, J. H. Stack, S. D. Emr, Division of Cellular and Molecular Medicine, University of California at San Diego, School of Medicine, and Howard Hughes Medical Institute, La Jolla, CA 92093. M. J. Fry and M. D. Waterfield, Ludwig Institute for Cancer Research, London, W1P 8BT, United Kingdom.

*To whom correspondence should be addressed.

*VPS*34 gene deleted (Δ*vps*34), and cells that overproduce Vps34p (Fig. 2A). The crude membrane fraction sedimented at 100,000*g* (P100) from wild-type yeast cells contained 90% of the cellular PI 3-kinase activity, whereas the supernatant (S100) contained only 10%. High amounts of PI 4-kinase activity, which phosphorylates the 4 position of the inositol ring (*23*), were present in both cellular fractions. The fact that 90% of the PI 3-kinase activity was present in the crude membrane pellet (P100), whereas only 50% of the Vps34 protein was recovered in this pellet, indicates that membrane-associated Vps34p has a higher specific activity (Fig. 2B). PI 3-kinase activity was not detected in either the P100 and S100 fractions isolated from the Δ*vps*34 strain. PI 4-kinase activity, however, was found in both the P100 and S100 fractions of the *vps*34 null mutant strain (Fig. 2A).

Cell lysates from yeast cells containing the *VPS*34 gene on a multicopy yeast plasmid contained fourfold more PI 3-kinase activity than was found in an equivalent amount of extract from wild-type cells. Unlike wild-type cells, however, the lipid kinase activity in extracts from the strain overproducing Vps34p was distributed nearly equally between the particulate (P100, 55%) and the soluble (S100, 45%) subcellular fractions (Fig. 2A). Approximately 20-fold more Vps34p was produced in the strain carrying the *VPS*34 gene on a multicopy number plasmid than in wild-type cells. The fact that most (95%) of Vps34p was found in the S100 fraction indicates that there are a limited number of sites for association of Vps34p with the crude membrane fraction or that a modification of Vps34p is required for the association and that the modifying system is saturated when Vps34p is overproduced (Fig. 2B). The association of Vps34p with the membrane or with one or more proteins in the membrane appears to result in the activation of the Vps34 PI 3-kinase activity (Fig. 2A).

To verify that Δ*vps*34 cells lack PI 3-kinase activity, we labeled the wild-type and Δ*vps*34 strains in vivo with [^{3}H]*myo*-inositol. The labeled lipids were isolated, deacylated, and separated by high-performance liquid chromatography (HPLC). Wild-type yeast contained two compounds that eluted at the same positions as the glycerophosphatidylinositol 3-phosphate (gPI3P) and glycerophosphatidylinositol 4-phosphate (gPI4P) standards, respectively, whereas only gPI4P was found in the lipids recovered from the Δ*vps*34 strain (Fig. 3). No multiply phosphorylated phosphoinositides [PI(3,4)P_2 or PI(3,4,5)P_3] were found under these conditions.

We obtained further evidence that the *VPS*34 gene encodes PI 3-kinase by demonstrating that PI 3-kinase activity can be specifically immunoprecipitated from cell extracts with antiserum that recognizes Vps34p. Protein A–Sepharose beads (Pharmacia) to which the antiserum to Vps34p was adsorbed were incubated with the S100 fraction isolated from the strain overproducing Vps34p. The resulting immune precipitate contained the PI 3-kinase activity (Fig. 4); no PI 3-kinase activity was detected in the supernatant fraction after the immune precipitation. By contrast, no PI

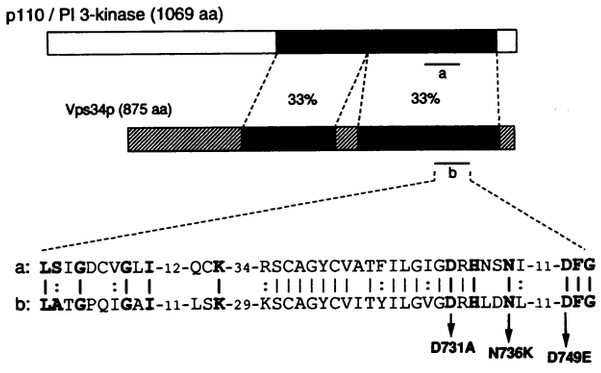

Fig. 1. Sequence similarity between Vps34p and the p110 subunit of the bovine brain PI 3-kinase. Black boxes correspond to regions of these proteins that share significant sequence similarity (33% amino acid identity). Shown below is a comparison of p110 (indicated by a) and Vps34p (indicated by b) sequences that contain the conserved sequence motifs present in the catalytic domains of protein kinases (bold letters) (*25, 26*). Single amino acid changes that were constructed by site-directed mutagenesis (*32*) of the *VPS*34 gene are also shown. In (a) and (b), vertical bars indicate identical amino acid residues, colons indicate conservative amino acid substitutions, and the numbers indicate amino acid spacing between conserved sequence elements. Abbreviations for the amino acids (aa) are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

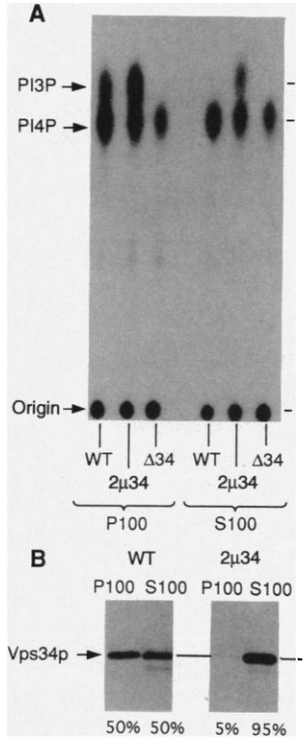

Fig. 2. PI 3-kinase activity of Vps34p and its subcellular localization. (**A**) Thin-layer chromatography (TLC) of PI kinase assay reaction products performed with the P100 and S100 fractions isolated from wild-type strain SEY6210 (WT), Δ*vps*34 strain PHY102 (Δ34), and strain SEY6210 bearing the *VPS*34 gene on a multicopy plasmid (2μ34) (*15*). Yeast cells were grown in yeast nitrogen base (YNB) medium (*33*), enzymatically converted into spheroplasts, lysed, and centrifuged at 500*g* for 5 min. The cleared cell extract was centrifuged at 100,000*g* for 30 min at 4°C. Protein (1 μg) from the pellet (P100) and supernatant (S100) fractions was assayed for PI kinase activity in the presence of [γ-^{32}P]ATP in a volume of 50 μl (*34*). Assays were incubated for 5 min at 25°C. Samples were separated on silica TLC plates (Merck) with the borate system (*35*). We determined the distribution of the total cellular PI 3-kinase activity between the subcellular fractions after normalizing for protein recoveries in the supernatant (70% of cellular protein) and the pellet (30% of cellular protein). For the wild-type strain, this resulted in a distribution of 90% of the PI 3-kinase activity in the P100 and 10% in the S100 fraction. For the strain overproducing Vps34p (2μ34), 55% of the PI 3-kinase activity was found in the P100 and 45% in the S100 fraction. We added sufficient exogenous substrate (10 μg of PI; Sigma) to attain linear reaction kinetics through a 10-min incubation period. No significant PI3P-phosphatase activity was detected under the assay conditions used. The reduced amount of PI4P produced in the Δ*vps*34 strain extract is most likely due to the slow growth of the mutant strain. A Phosphor-Imager (Molecular Dynamics) was used for quantitation. (**B**) Distribution of Vps34p between the P100 and S100 fractions of SEY6210 (WT) and the strain overproducing Vps34p (2μ34). Spheroplasts were labeled in selective minimal media in the presence of Tran-^{35}S-label (10 μCi/$OD_{600\ nm}$, ICN) for 30 min and incubated at 30°C for 60 min after the addition of methionine and cysteine to a final concentration of 25 mM (*15*). The spheroplasts were lysed and separated into P100 and S100 fractions as described in (A). We used antiserum to Vps34p to immunoprecipitate Vps34p (*15*), and the immunoprecipitated proteins were separated by SDS–polyacrylamide gel electrophoresis (PAGE) (8% gel). The percentage distribution between P100 and S100 fractions was deteremined by densitometry of the audioradiogram (LKB Ultroscan XL) and is indicated under each lane. Relative to the wild-type strain, one-tenth as much sample from the overproducing strain (2μ34) was loaded onto the gel.

4-kinase activity was detected in the pellet fraction. It remained in the supernatant (Fig. 4). Preimmune serum did not precipitate any PI 3-kinase or PI 4-kinase activity.

The regions of sequence similarity between Vps34p and p110 are in the COOH-terminal parts of the proteins (Fig. 1). The degree of identity in this region is 33%, and the degree of similarity is >57% if conservative amino acid changes are included [comparison of amino acids 315 to 820 of Vps34p with amino acids 539 to 1014 of p110 with gap analysis (*24*)]. The regions of similarity between Vps34p and p110 include motifs that are also highly conserved in the catalytic domains of protein kinases (Fig. 1) (*25*). Vps34p and p110, however, contain only a subset of the structural features conserved among protein kinases. The crystal structure of the adenosine 3′,5′-monophosphate (cAMP)–dependent protein kinase (PKA) indicates that these conserved motifs participate in adenosine trisphosphate (ATP) binding and phosphate transfer (*26*). These motifs are the glycine-rich region (GXGXXG; G is glycine and X represents any amino acid), which interacts with nontransferable phosphates; the catalytic loop region (DXHXXN; D is aspartic acid, H is histidine, and N is asparagine); and the DFG (F is phenylalanine) motif, which is the smallest and most highly conserved structural feature among protein kinases.

We used site-directed mutagenesis to construct three *vps*34 point mutations to study the functional significance of the sequences shared between Vps34p and protein kinases. Asp^{731} was replaced by alanine (D731A), Asn^{736} was changed to lysine (N736K), and Asp^{749} was replaced by glutamic acid (D749E) (Fig. 1). These *vps*34 mutant alleles were introduced into a Δ*vps*34 strain on a multicopy plasmid vector. No PI 3-kinase activity was detected in extracts from any of the three mutants (Fig. 5A). In vivo [^{35}S]methionine labeling experiments demonstrated that all of the mutants expressed Vps34p (Fig. 5B), and the mutant proteins exhibited stability similar to that of the wild type. The various *vps*34 point mutant constructs did not complement any of the phenotypes associated with a Δ*vps*34 mutation (*15*). All three strains were temperature-sensitive for growth and had a severe defect in vacuolar protein sorting. Unlike wild-type cells in which the Golgi apparatus–modified precursor form of CPY (p2CPY) is transported to the vacuole and matured (mCPY), all of the p2CPY was secreted by the mutant cells. No CPY was detected inside these cells (Fig. 5C).

The data presented in this study implicate PI 3-kinase activity in the regulation of vacuolar protein sorting. The precise function of PI3P or of multiply phosphorylated 3-phosphoinositides is not yet known. At least two general models can be considered for the role of phosphoinositides in protein sorting. (i) They may alter the biophysical properties of the lipid bilayer in the sorting compartment: the increase of the PI3P (a charged phospholipid) in the outer leaflet of the membrane could change the molecular organization of the outer leaflet, thus increasing the curvature of the membrane and thereby stimulating vesicle emergence, as has been postulated in the lipid-bilayer couple hypothesis (*27*). (ii)

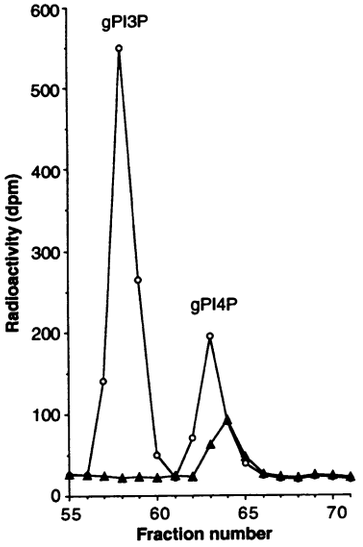

Fig. 3. In vivo PI kinase activity in wild-type and Δ*vps*34 strains. The HPLC elution profiles of tritiated glycerophosphoinositols isolated from SEY6210 (WT; open circles) and PHY102 (Δ*vps*34; closed triangles) are shown. Strains were grown in the presence of 30 μCi (6 μCi/ml) of [^{3}H]*myo*-inositol (18.8 Ci/mmol; Amersham) over six generations at 25°C in synthetic Wickerham's minimal growth medium without inositol (*36*). Vitamins and amino acids were added as required (Difco). Lipids were extracted with acidified chloroform–methanol and deacylated; the resulting glycerophosphoinositols were separated by HPLC (*37*). Fractions were analyzed in a liquid scintillation counter (Beckman LS1801). Tritiated standards of glycerophosphatidylinositol, gPI4P, and glycerophosphatidylinositol 4,5-bisphosphate were prepared from [^{3}H]PI, [^{3}H]PI4P and [^{3}H]PI(4,5)P_2, respectively, and we prepared a glycerophosphatidylinositol 3-monophosphate (gPI3P) standard by phosphorylating PI in vitro with purified bovine brain PI 3-kinase in the presence of [γ-^{32}P]ATP.

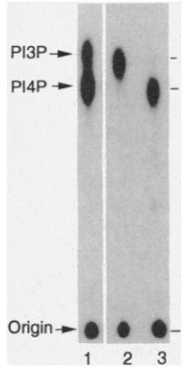

Fig. 4. Immunoprecipitation of PI 3-kinase activity as in Fig. 2A. PI kinase assays were done on the total S100 fraction of the Vps34p-overproducing strain (lane 1), the sedimented Vps34p immune complex (lane 2), and the supernatant from the immunoprecipitation (lane 3). Protein A–Sepharose beads were adsorbed to antibodies specific for Vps34p (*15*) and washed, and the supernatant fraction (S100) of SEY6210 containing *VPS*34 on a yeast multicopy plasmid (2μ34) was incubated with the beads for 4 hours at 4°C. PI kinase assays were done after we washed the immune complexes (four times with 0.5% Tween-20 in tris-buffered saline and six times with tris-buffered saline). After the immunoprecipitation reactions, 25 to 50% of the total PI kinase activities were recovered.

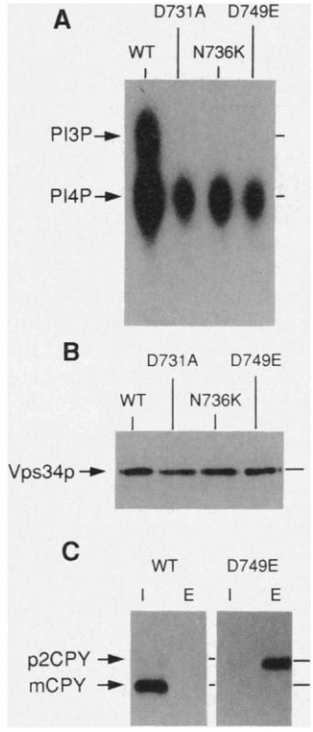

Fig. 5. The *vps*34 point mutations. (**A**) TLC of PI kinase assay reaction products done with the P100 fractions of wild-type strain SEY6210 (WT) and *vps*34 point mutants. All *vps*34 mutant alleles were present on multicopy plasmids in PHY102 (Δ*vps*34). (**B**) Expression of wild-type and mutant *VPS*34 proteins. Cells were labeled with Tran-^{35}S-label for 20 min at 30°C. Proteins precipitated in trichloracetic acid were resuspended and incubated with antiserum to Vps34p; proteins in the immunoprecipitate were separated by SDS-PAGE (10% gel) as described (*15*). (**C**) Sorting of CPY in wild-type yeast and in one of the point mutants (D749E). Nearly identical results were obtained with the D731A and N736K mutants. Spheroplasts were grown in selective minimal media in the presence of Tran-^{35}S-label for 5 min and chased for 30 min. The labeled cultures were centrifuged for 2 min at 13,000*g* and separated into sedimented (I, intracellular) and supernatant (E, extracellular) fractions. The amount of CPY in each fraction was assessed by quantitative immunoprecipitation with CPY-specific antisera, followed by SDS-PAGE as described (*31*).

Phosphoinositides could serve as specific membrane targets that bind proteins required for the formation of transport vesicles, such as the polypeptides of the adaptor complex that link clathrin to the cytoplasmic tails of certain transmembrane receptor proteins (for example, the mannose-6-phosphate receptor) (*28*). Vesicles from rat adipocytes that contain the glucose transporter also contain PI 4-kinase, which may regulate the transport (fusion) of these vesicles with the plasma membrane in response to insulin (*29*). In addition to its potential role in signaling cell proliferation, PI 3-kinase associated with receptor protein tyrosine kinases at the plasma membrane may also take part in the endocytosis and down-regulation (lysosomal degradation) of these receptors. In this way, the duration and the magnitude of the growth signal might be modulated. The association of Vps34p with the membrane appears to be mediated by the product of another *VPS* gene, *VPS15*. The *VPS15* gene encodes a membrane-associated protein kinase (Vps15) (*30, 31*) that can be chemically cross-linked to Vps34p (*21*). This raises the possibility that the Vps15 and Vps34 proteins may function together as components of a signal transduction complex that regulates intracellular protein sorting decisions.

REFERENCES AND NOTES

1. L. C. Cantley *et al.*, *Cell* **64**, 281 (1991).
2. J. M. Backer *et al.*, *EMBO J.* **11**, 3469 (1992).
3. S. Soltoff, S. Rabin, L. Cantley, D. Kaplan, *J. Biol. Chem.* **267**, 17472 (1992).
4. M. J. Berridge and R. F. Irvine, *Nature* **341**, 197 (1989).
5. M. Whitman *et al.*, *ibid.* **332**, 644 (1988).
6. D. L. Lips *et al.*, *J. Biol. Chem.* **264**, 8759 (1989).
7. L. A. Serunian *et al.*, *ibid.*, p. 17809.
8. C. Carpenter *et al.*, *ibid.* **265**, 19704 (1990).
9. F. Shibasaki, Y. Homma, T. Takenawa, *ibid.* **266**, 8108 (1991).
10. M. Otsu *et al.*, *Cell* **65**, 91 (1991).
11. I. D. Hiles *et al.*, *ibid.* **70**, 419 (1992).
12. C. A. Koch, D. Anderson, M. F. Moran, C. Ellis, T. Pawson, *Science* **252**, 668 (1991).
13. P. Hu *et al.*, *Mol. Cell. Biol.* **12**, 981 (1992).
14. C. J. McGlade *et al.*, *ibid.*, p. 991.
15. P. K. Herman and S. D. Emr, *ibid.* **10**, 6742 (1990).
16. S. Kornfeld and I. Mellman, *Annu. Rev. Cell Biol.* **5**, 483 (1989).
17. D. J. Klionsky, P. K. Herman, S. D. Emr, *Microbiol. Rev.* **54**, 266 (1990).
18. J. E. Rothman and L. Orci, *Nature* **355**, 409 (1992).
19. T. Stevens *et al.*, *Cell* **30**, 439 (1982).
20. T. R. Graham and S. D. Emr, *J. Cell Biol.* **114**, 207 (1991).
21. J. H. Stack, P. K. Herman, P. V. Schu, S. D. Emr, *EMBO J.*, in press.
22. K. R. Auger, C. L. Carpenter, L. C. Cantley, L. Varticovski, *J. Biol. Chem.* **264**, 20181 (1989).
23. G. Endemann, S. N. Dunn, L. C. Cantley, *Biochemistry* **26**, 6845 (1987).
24. Genetics Computer Group sequence analysis package; University of Wisconsin.
25. S. K. Hanks *et al.*, *Science* **241**, 42 (1988).
26. D. R. Knighton *et al.*, *ibid.* **253**, 407 (1991).
27. M. P. Sheetz and S. J. Singer, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4457 (1974).
28. B. M. Pearse and M. S. Robinson, *Annu. Rev. Cell Biol.* **6**, 151 (1990).
29. R. L. Del Vecchio and P. F. Pilch, *J. Biol. Chem.* **266**, 13278 (1991).
30. P. K. Herman *et al.*, *Cell* **64**, 425 (1991).
31. P. K. Herman, J. H. Stack, S. D. Emr, *EMBO J.* **10**, 4049 (1991).
32. T. Kunkel, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5463 (1985).
33. F. Sherman, G. R. Fink, L. W. Lawrence, *Methods in Yeast Genetics: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1979).
34. M. Whitman *et al.*, *Nature* **315**, 239 (1985).
35. J. P. Walsh, K. K. Caldwell, P. W. Majerus, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9184 (1991).
36. L. J. Wickerham, *J. Bacteriol.* **52**, 293 (1946).
37. K. R. Auger *et al.*, *Cell* **57**, 167 (1989).

38. We thank members of the Emr lab for discussions and G. Huyer for his help in optimizing the PI 3-kinase assay. Supported by the Howard Hughes Medical Institute, a grant from the NSF (to S.D.E.), and by a fellowship from the Deutsche Forschungsgemeinschaft (to P.V.S.). S.D.E. is an investigator of the Howard Hughes Medical Institute.

8 October 1992; accepted 13 January 1993

Natural Selection and the Origin of *jingwei*, a Chimeric Processed Functional Gene in *Drosophila*

Manyuan Long* and Charles H. Langley

The origin of new genes includes both the initial molecular events and subsequent population dynamics. A processed *Drosophila* alcohol dehydrogenase (*Adh*) gene, previously thought to be a pseudogene, provided an opportunity to examine the two phases of the origin of a new gene. The sequence of the processed *Adh* messenger RNA became part of a new functional gene by capturing several upstream exons and introns of an unrelated gene. This novel chimeric gene, *jingwei*, differs from its parent *Adh* gene in both its pattern of expression and rate of molecular evolution. Natural selection participated in the origin and subsequent evolution of this gene.

How genes with novel functions evolve remains a fundamental and fascinating question. Gene duplications (*1*), exon shuffling, and processed genes (*2*) have been suggested as important sources of novel genes, but little is known about the evolutionary mechanisms or the participation of natural selection in their early history. Our analysis of the structure, expression, and evolution of a putative processed *Adh* pseudogene in *Drosophila* provided an opportunity to examine both the early molecular events and the evolutionary processes that created it. The *jingwei* (*jgw*) gene, as we named it (*3*), is a locus located on chromosome 3 in the *Drosophila* sibling species *D. teissieri* and *D. yakuba*. A part of *jgw* was initially observed to hybridize to the *Adh* probe (*4*). Further analysis suggested that in a single event, in the ancestor of the two species the *Adh* portion of this potential pseudogene was retrotransposed from an mRNA of the *Adh* locus on chromosome 2 (*5*). To understand the molecular population genetics of this potential pseudogene, we investigated its within-species DNA sequence variation. However, our results convince us that *jgw* is not a pseudogene and that the *Adh*-derived sequence is a part of a novel gene.

The *Adh*-derived portion was sequenced from ten *jgw* alleles of *D. teissieri* and of 20 *jgw* alleles of *D. yakuba* collected from natural populations (*6*). Figure 1 shows the distribution of nucleotide polymorphisms within species and the variation between species in a 765-bp segment that corresponds to the protein coding region of the *Adh* gene. Only one possible ancestral polymorphism was apparent (site 782). A summary of the DNA sequence variation in this region is presented (Table 1).

Unexpectedly, most polymorphisms were silent (eight out of ten in *D. teissieri* and 19 out of 21 in *D. yakuba*). If *jgw* were a pseudogene in which mutations had no phenotypic effect, most changes would be at replacement sites and there would be a reasonable frequency of stop codons (*7*). No new stop codons were present. Another prediction of the pseudogene hypothesis is that a pseudogene has a higher overall level of nucleotide variation. Estimates of DNA sequence polymorphism in *jgw* were similar to that found in the *Adh* gene (Table 1) and are typical of many functional genes in *Drosophila* (≈0.005) (*8*). A comparison of between-species divergence also revealed a significant bias toward silent versus replacement substitutions, and the degree of bias was smaller than for the within-species comparison. The bias was noted in the earlier study (*5*) but by itself was not large enough to convince the authors that this was not a pseudogene. In addition, insertions and deletions are abundant in the evolution of mammalian pseudogenes (*7*). In *jgw*, no length polymorphism or divergence was observed in the coding region (*9*). These molecular population genetic observations imply that the *Adh*-derived sequence is all or part of *jgw*, a new functional gene in *D. teissieri* and *D. yakuba*. This conclusion motivated our search

Section of Genetics, Section of Evolution and Ecology, and Center for Population Biology, University of California at Davis, Davis, CA 95616.

*To whom correspondence should be addressed.