CG6051 PI3P Phosphatase Activity Hypothesis: Deep Research Report

Summary

The hypothesis that *Drosophila melanogaster* CG6051 (UniProt Q9VB70) has phosphatidylinositol-3-phosphate (PI3P) phosphatase activity (GO:0004438) is **refuted — this is an over-annotation**. CG6051 encodes "Lateral signaling target protein 2 homolog" (LST2_DROME), a FYVE zinc finger domain-containing adaptor protein of the lst-2/ANKFY1 family. The protein contains a well-characterized PI3P-*binding* domain (FYVE, IPR000306) but completely lacks any phosphatase catalytic domain — no myotubularin domain, no protein tyrosine phosphatase (PTP) domain, no dual-specificity phosphatase (DSP) domain, and critically, no CX5R active site motif that is the hallmark of PI phosphatases. The annotation derives exclusively from automatic electronic annotation (IEA via UniProtKB-ARBA, GO_REF:0000117), with no supporting experimental, phylogenetic, or curated evidence.

Cross-species ortholog analysis provides strong negative evidence against this annotation. The *C. elegans* ortholog LST-2 (Q9TZD0) is annotated with GO:0032266 "phosphatidylinositol-3-phosphate binding" (ISS evidence) and carries no phosphatase terms. The human ortholog ANKFY1 (Q9P2R3) is annotated with GO:1901981 "phosphatidylinositol phosphate binding" (IDA evidence — direct experimental assay) and similarly has no phosphatase annotations. CG6051 is the sole member of the lst-2 family in any organism that carries phosphatase GO terms, and all such terms are IEA-only. Meanwhile, the genuine *Drosophila* PI3P phosphatases — Mtm (myotubularin), EDTP (egg-derived tyrosine phosphatase/MTMR14), and dMtmr6 (CG3530) — all possess distinct myotubularin-type catalytic domains with experimentally validated phosphatase activity (IDA evidence).

The recommended curation action is to **remove** GO:0004438 and related phosphatase annotations from CG6051, and to **consider adding** GO:0032266 (PI3P binding) to accurately capture the FYVE domain's function. The two existing IBA annotations — GO:0042059 (negative regulation of EGFR signaling pathway) and GO:0031901 (early endosome membrane) — are consistent with the protein's adaptor function and should be retained.

Executive Judgment

Verdict: OVER-ANNOTATED (refuted)

The hypothesis that CG6051 has phosphatidylinositol-3-phosphate phosphatase activity (GO:0004438) is **not supported** by domain architecture, sequence analysis, ortholog comparison, or any published experimental evidence. The GO:0004438 annotation is an artifact of automatic rule-based annotation (ARBA) that has inappropriately conflated PI3P *binding* (a function of the FYVE domain) with PI3P *catalytic dephosphorylation* (a function requiring a phosphatase domain that CG6051 does not possess).

Most important caveat: No direct biochemical assay of CG6051 phosphatase activity has been published (positive or negative), so the absence of phosphatase activity is inferred from domain architecture and ortholog comparison rather than from a direct experimental test. However, the complete absence of any recognizable phosphatase catalytic domain makes intrinsic phosphatase activity biologically implausible.

Key Findings

Finding 1: CG6051 lacks any phosphatase catalytic domain — the PI3P phosphatase annotation is an over-annotation

Domain analysis of CG6051 (989 amino acids) reveals that the protein contains only a FYVE zinc finger domain (InterPro IPR000306, Pfam PF01363) spanning approximately residues 915–964, with eight conserved zinc-coordinating cysteine residues characteristic of the FYVE fold. This domain is well-established as a PI3P-*binding* module — it recognizes the head group of phosphatidylinositol-3-phosphate on endosomal membranes and tethers the protein to these compartments. Crucially, the FYVE domain has no catalytic activity; it is a lipid-recognition domain, not an enzyme.

A comprehensive search for phosphatase catalytic motifs across the full CG6051 sequence found no myotubularin phosphatase domain, no PTP domain, and no CX5R active site motif. The CX5R motif (Cys-X5-Arg) is the hallmark catalytic signature of the PTP superfamily, including all myotubularin-related PI3P phosphatases. As established in structural reviews of PI phosphatase architecture, "PI phosphatases are a large collection of enzymes that are evolved from at least two disparate ancestors. One group is distantly related to endonucleases,

which apply divalent metal ions for phosphoryl transfer. The other group is related to protein tyrosine phosphatases, which contain a highly conserved active site motif Cys-X5-Arg (CX5R)" (PMID: 25264170). CG6051 lacks both types of catalytic architecture. The only CX5R-like sequences found in CG6051 map to the zinc-coordinating cysteines of the FYVE domain — structural residues with no catalytic role.

Ortholog comparison reinforces this conclusion. The *C. elegans* ortholog LST-2 (Q9TZD0) is annotated with GO:0032266 "phosphatidylinositol-3-phosphate binding" with ISS evidence, and carries no phosphatase GO terms whatsoever. The human ortholog ANKFY1 (Q9P2R3) — the best-characterized member of this protein family — is annotated with GO:1901981 "phosphatidylinositol phosphate binding" based on direct experimental assay (IDA evidence), and also carries no phosphatase annotations. ANKFY1 has been functionally characterized as a Rab5 effector involved in endosome fusion and macropinocytosis — an adaptor protein, not an enzyme. CG6051/Q9VB70 is the **only** lst-2 family member in any organism that carries phosphatase GO terms, and all three of its phosphatase annotations (GO:0004438, GO:0052629, GO:0004721) derive from IEA:UniProtKB-ARBA — automatic rules with no manual curation or experimental validation.

Finding 2: IBA annotations support adaptor/EGFR regulation function, not phosphatase activity

Analysis of all GO annotations for Q9VB70 in QuickGO reveals 13 total annotations. Of these, 10 are IEA (automatic, GO_REF:0000117/ARBA), 2 are IBA (phylogenetic inference from GO_Central, GO_REF:0000033), and 1 is IEA from InterPro (GO_REF:0000002). The two phylogenetically-inferred IBA annotations are:

1. **GO:0042059** — negative regulation of epidermal growth factor receptor signaling pathway
2. **GO:0031901** — early endosome membrane

These annotations, derived from multi-species phylogenetic analysis by the GO Consortium (GO_Central), are consistent with the known adaptor function of the lst-2 family: binding PI3P at endosomal membranes to regulate EGFR trafficking and signaling. Notably, the IBA pipeline did **not** propagate any phosphatase activity annotation to CG6051, further confirming that the phosphatase assignment is not supported by phylogenetic evidence. The fact that a rigorous, phylogeny-aware annotation pipeline arrived at adaptor/regulatory function rather than catalytic phosphatase activity is significant negative evidence against GO:0004438.

Mechanistic Model / Interpretation

What CG6051 actually does: an endosomal adaptor protein

Based on the convergent evidence from domain architecture, ortholog function, and phylogenetic annotation, CG6051/LST2 functions as a **FYVE domain-containing endosomal adaptor protein** that:

1. **Binds PI3P** via its C-terminal FYVE zinc finger domain, targeting it to the cytoplasmic face of early endosome membranes
2. **Negatively regulates EGFR signaling** (IBA annotation), likely by promoting receptor sorting through the endosomal pathway toward lysosomal degradation
3. **Acts as an adaptor/scaffold** (IEA:ARBA, but supported by ortholog biology), recruiting downstream effectors to endosomal compartments

```
Functional Model of CG6051/LST2:

  Plasma membrane    EGFR-ligand complex
       |                    |
       v                    v
  [Endocytosis] ---------> [Early endosome]
                                |
                           PI3P in membrane
                                |
                     CG6051 FYVE domain BINDS PI3P
                     (NOT dephosphorylation)
                                |
                     CG6051 adaptor scaffolding
                                |
                     Promotes EGFR sorting/degradation
                                |
                     Negative regulation of EGFR signaling

  Contrast with TRUE PI3P phosphatases:

  Mtm / EDTP / dMtmr6:
    [Myotubularin phosphatase domain] + [CX5R active site]
         |
    PI3P -------> PI  (catalytic dephosphorylation)
         |
    Regulates endosome maturation, autophagy
```

The critical distinction: binding vs. catalysis

The FYVE domain is one of several well-characterized PI3P-binding modules (others include the PX domain and some PH domains). As established by studies of FYVE domain-containing proteins such as EEA1 (PMID: 34555023), Phafin1/2 (PMID: 37175801, PMID: 38947768), SPG15 (PMID: 33464297), and plant FYVE4 (PMID: 33772801), FYVE domains serve as membrane-targeting modules that recruit their host proteins to PI3P-enriched endosomal surfaces. None of these FYVE-only proteins are annotated as PI3P phosphatases. The FYVE domain reads the lipid signal; it does not destroy it.

True *Drosophila* PI3P phosphatases belong to the myotubularin family and are functionally distinct:

Protein	Gene	Domain architecture	GO:0004438 evidence	Function
Mtm	CG9115	Myotubularin PTP domain + PH-GRAM	IDA (FlyBase)	PI3P phosphatase, Hippo pathway, actomyosin regulation
EDTP	CG6016	Myotubularin PTP domain	IDA/IMP	PI3P phosphatase, autophagy suppressor, aging
dMtmr6	CG3530	Myotubularin PTP domain	IDA	PI3P 3-phosphatase, autophagic flux regulation
CG6051	**CG6051**	**FYVE only (no PTP)**	**IEA only**	**PI3P BINDING adaptor**

Evidence Base

Primary literature supporting the refutation

Hsu & Bhatt-Grover, 2015 — *The structure of phosphoinositide phosphatases: Insights into substrate specificity and catalysis* (PMID: 25264170) This structural review establishes that all PI phosphatases of the PTP-related group "contain a highly conserved active site motif Cys-X5-Arg (CX5R)." CG6051 lacks this motif entirely, ruling out classification as a PI phosphatase of this

superfamily. This paper provides the structural framework for understanding why the FYVE domain alone is insufficient for phosphatase activity. **Confidence: High** — well-established structural biology, directly relevant.

Hu et al., 2022 — *Myotubularin functions through actomyosin to interact with the Hippo pathway* (PMID: 36285521) This study of *Drosophila* Mtm demonstrates that the true PI3P phosphatase in *Drosophila* is Mtm, which "regulates membrane phospholipid PI(3)P dynamics" through its myotubularin phosphatase domain. The paper shows the functional role of a bona fide PI3P phosphatase in growth control, providing a clear comparator to CG6051. **Confidence: High** — direct functional study of the true Drosophila PI3P phosphatase.

Allen et al., 2020 — *A conserved myotubularin-related phosphatase regulates autophagy by maintaining autophagic flux* (PMID: 32915229) Identified CG3530/dMtmr6 as another *Drosophila* PI3P 3-phosphatase of the MTMR6 family, demonstrating it "functions as a regulator of autophagic flux in multiple Drosophila cell types." This confirms that *Drosophila* has multiple genuine myotubularin-family PI3P phosphatases, all with proper catalytic domains — CG6051 is not among them. **Confidence: High** — functional screen with validation.

Varga et al., 2023 — *A conserved MTMR lipid phosphatase increasingly suppresses autophagy in brain neurons during aging* (PMID: 36528685) Demonstrates that EDTP (the *Drosophila* MTMR14 ortholog) functions as an autophagy repressor through its lipid phosphatase activity. Provides additional context for the myotubularin-family phosphatases in *Drosophila*. **Confidence: High** — direct in vivo study.

Kovacs et al., 2015 — *AUTEN-67, an autophagy-enhancing drug candidate with potent antiaging and neuroprotective effects* (PMID: 26312549) Identified AUTEN-67 as "specific inhibitors of MTMR14, a myotubularin-related phosphatase antagonizing the formation of autophagic membrane structures." This study explicitly targets EDTP/MTMR14 as a phosphatase, further delineating which *Drosophila* proteins genuinely have PI3P phosphatase activity. **Confidence: High** — drug screen against a validated phosphatase target.

Literature on FYVE domain biology (supporting binding, not catalysis)

Kutateladze, 2021 — *Membrane-binding mechanism of the EEA1 FYVE domain revealed by multi-scale molecular dynamics simulations* (PMID: 34555023) Characterizes the FYVE domain as a PI3P-binding module with "specific PI3P-interactions" providing ~50-60 kJ/mol binding energy. The study shows that FYVE domains bind and recognize PI3P but do not catalyze its dephosphorylation. **Confidence: High** — detailed structural/biophysical study.

Schink et al., 2023 — *Phafins Are More Than Phosphoinositide-Binding Proteins* (PMID: 37175801) Reviews Phafin proteins containing PH+FYVE domains as "PtdIns(3)P effectors" that act as "adaptor proteins that recruit other binding partners." Despite containing FYVE domains, Phafin proteins are classified as PI3P-binding adaptors, not phosphatases. This supports the paradigm that FYVE domain = binding, not catalysis. **Confidence: High** — comprehensive review.

Hori et al., 2021 — *Rag GTPases and phosphatidylinositol 3-phosphate mediate recruitment of the AP-5/SPG11/SPG15 complex* (PMID: 33464297) Shows that SPG15 uses its FYVE domain for "PI3P binding" to recruit the AP-5/SPG11/SPG15 complex to endosomes — another example of FYVE as a targeting domain, not an enzyme. **Confidence: High** — functional cell biology.

Evidence Matrix

#	Citation	Evidence Type	Direction	Claim Tested	Key Finding	Context	Confi
1	UniProt Q9VB70, InterPro	Computational / Database	Refutes	CG6051 has PI3P phosphatase activity	FYVE domain only; no phosphatase domain of any type	*D. melanogaster*, domain annotation	High
2	Sequence motif analysis	Computational	Refutes	CG6051 has a phosphatase active site	No CX5R motif, no HCxxGxxR(S/T) loop, no WPD loop	Full 989-aa sequence scan	High
3	PMID: 25264170	Structural review	Qualifies	PI phosphatases require CX5R	CX5R is a conserved hallmark; CG6051 lacks it	PI phosphatase superfamily	High
4	UniProt Q9TZD0 (*Ce* LST-2)	Ortholog / Computational	Refutes	lst-2 family has phosphatase activity	Annotated PI3P binding (ISS), no phosphatase terms	*C. elegans* ortholog	Mode
5	UniProt Q9P2R3 (*Hs* ANKFY1)	Ortholog / Direct assay	Refutes	lst-2 family has phosphatase activity	Annotated PIP binding (IDA), no phosphatase terms	Human ortholog	High
6	PMID: 36285521	Direct assay / Mutant	Qualifies	Drosophila Mtm is the true PI3P phosphatase	Mtm dephosphorylates PI3P via myotubularin domain	*D. melanogaster* wing disc	High
7	PMID: 32915229	Direct assay / Screen	Qualifies	dMtmr6/CG3530 is a distinct PI3P phosphatase	CG3530 is an MTMR6-family 3-phosphatase	*D. melanogaster*, multiple cell types	High
8	QuickGO Q9VB70	Database	Refutes	Non-IEA evidence	All phosphatase terms are IEA; IBA terms	GO annotation database	High

#	Citation	Evidence Type	Direction	Claim Tested	Key Finding	Context	Confi
				supports phosphatase	support adaptor function		
9	PMID: 34555023	Structural / Biophysical	Qualifies	FYVE domains bind but don't hydrolyze PI3P	FYVE binds PI3P with ~50-60 kJ/mol; no catalysis	EEA1 FYVE domain, MD simulation	High
10	PMID: 37175801	Review	Qualifies	FYVE proteins are PI3P effectors, not enzymes	Phafin FYVE proteins act as "adaptor proteins"	Multiple organisms	High

GO Curation Implications

Recommended Curation Actions (leads requiring curator verification)

Remove (3 MF terms + 1 BP term):

1. **GO:0004438** (phosphatidylinositol-3-phosphate phosphatase activity, MF) — **Remove**. IEA:ARBA only. No phosphatase catalytic domain. Contradicted by ortholog annotations.
2. **GO:0052629** (phosphatidylinositol-3,5-bisphosphate 3-phosphatase activity, MF) — **Remove**. Same rationale as above.
3. **GO:0004721** (phosphoprotein phosphatase activity, MF) — **Remove**. No phosphatase domain of any kind.
4. **GO:0046856** (phosphatidylinositol dephosphorylation, BP) — **Remove**. Process term inappropriately assigned without underlying catalytic function.

Retain (2 IBA terms):

1. **GO:0042059** (negative regulation of EGFR signaling pathway, BP) — **Retain**. IBA from GO_Central; consistent with lst-2 family function.

2. **GO:0031901** (early endosome membrane, CC) — **Retain**. IBA from GO_Central; consistent with FYVE-PI3P targeting.

Consider adding (1 MF term):

1. **GO:0032266** (phosphatidylinositol-3-phosphate binding, MF) — **Add** with ISS evidence, citing *C. elegans* LST-2 (Q9TZD0) and human ANKFY1 (Q9P2R3). This accurately captures the FYVE domain's function. Alternatively, the more general GO:1901981 (phosphatidylinositol phosphate binding) could be used to match the human IDA annotation.

Review (1 MF term):

1. **GO:0060090** (molecular adaptor activity, MF) — Currently IEA:ARBA, but supported by ortholog biology (ANKFY1 is a characterized Rab5 effector/adaptor). Could be upgraded if direct Drosophila evidence becomes available.

Summary of GO term changes

Current Term	Current Evidence	Action	Replacement
GO:0004438 PI3P phosphatase activity	IEA:ARBA	Remove	GO:0032266 PI3P binding
GO:0052629 PI(3,5)P2 3-phosphatase	IEA:ARBA	Remove	—
GO:0004721 phosphoprotein phosphatase	IEA:ARBA	Remove	—
GO:0046856 PI dephosphorylation	IEA:ARBA	Remove	—
GO:0042059 neg reg EGFR signaling	IBA	Retain	—
GO:0031901 early endosome membrane	IBA	Retain	—

Conflicts and Alternatives

Source of the over-annotation

The GO:0004438 annotation derives from **UniProtKB-ARBA** (Automatic Rule-Based Annotation, GO_REF:0000117). ARBA rules propagate annotations from well-characterized proteins to related entries based on sequence features. The most likely source of error is that the ARBA system has associated CG6051 with a rule derived from proteins that possess both FYVE domains **and** phosphatase domains. Some MTMR family members (e.g., MTMR3, MTMR4) contain FYVE-like domains alongside their catalytic phosphatase domains. If the ARBA rule was built on features from such dual-domain proteins and then triggered by the FYVE domain alone, it would produce exactly this type of over-annotation — attributing catalytic activity to a protein that has only the binding/targeting domain.

No paralog confusion

CG6051 is not a paralog of any known *Drosophila* phosphatase. The myotubularin family (Mtm/CG9115, EDTP/CG6016, dMtmr6/CG3530) is phylogenetically distinct from the lst-2/ANKFY1 family. There is no basis for transferring phosphatase annotations between these families.

Organism-specific considerations

Drosophila CG6051 is the **only** lst-2 family member in any organism annotated with phosphatase GO terms. Both the *C. elegans* (Q9TZD0) and human (Q9P2R3) orthologs consistently annotate PI3P interaction as *binding*, not catalysis. This cross-species consistency argues strongly that the phosphatase annotation on CG6051 is an error specific to the ARBA rule system, not a reflection of organism-specific neofunctionalization.

MTMR4 as a potential source of confusion

One paper in the literature describes MTMR4 as "a FYVE domain-containing dual-specificity protein phosphatase (DUSP)" that can "interact with BMP/Dpp signaling" in *Drosophila* (PMID: 23150675). However, MTMR4 (human) and its *Drosophila* homolog CG3632 contain both a FYVE domain **and** a myotubularin-type phosphatase domain — they are dual-domain proteins. CG6051 has only the FYVE domain. This distinction is critical: the FYVE domain targets the protein to PI3P-containing membranes, while the phosphatase domain (when present) catalyzes dephosphorylation.

Knowledge Gaps

Gap	What was checked	Why it matters	What would resolve it
No direct biochemical assay of CG6051	PubMed search for CG6051 phosphatase activity — no papers found	A direct in vitro phosphatase assay would definitively confirm or refute the annotation	Recombinant CG6051 phosphatase activity assay against PI3P substrate
ARBA rule identity unknown	InterPro2GO mappings checked (no phosphatase terms for FYVE); specific ARBA rule ID not publicly queryable	Understanding which ARBA rule generated this annotation could prevent similar errors for other proteins	UniProt could expose ARBA rule provenance in annotation details
No published functional study of CG6051 in *Drosophila*	PubMed searches for "CG6051," "lateral signaling target Drosophila" — no primary research papers found	Even the adaptor function is inferred from orthologs, not direct *Drosophila* experiments	Genetic or cell-biological characterization of CG6051 in *Drosophila*
CG6051 expression and tissue localization	Not systematically checked (no programmatic access to FlyBase expression data in this session)	Could support or refute endosomal localization in specific *Drosophila* tissues	FlyBase modENCODE expression data or antibody staining experiments
Whether CG6051 has any enzymatic activity at all	No structural prediction or active site analysis beyond motif search	Remote possibility of a non-canonical enzymatic activity	AlphaFold structure analysis, active site prediction tools

Proposed Follow-up Experiments / Actions

Priority 1: Curation action (no experiment needed)

- **Remove GO:0004438, GO:0052629, GO:0004721, and GO:0046856** from CG6051/Q9VB70 based on the domain architecture and ortholog evidence presented in this report. This is a database correction, not a scientific question.
- **Add GO:0032266** (PI3P binding) with ISS evidence, citing *C. elegans* LST-2 (Q9TZD0) as the basis for sequence similarity transfer.
- **Flag ARBA rule** for UniProt review — the rule generating phosphatase annotations for lst-2 family members is producing false annotations.

Priority 2: Definitive experimental test

- **In vitro phosphatase assay:** Express and purify recombinant CG6051 (full-length or FYVE domain-containing fragment) and test for phosphatase activity against PI3P, PI(3,5)P2, and generic phosphatase substrates (e.g., pNPP, DiFMUP). Use purified *Drosophila* Mtm as positive control and heat-inactivated CG6051 as negative control. This is the single most definitive experiment but is low priority given the overwhelming domain evidence.

Priority 3: Functional characterization

- **PI3P binding assay:** Perform lipid overlay (PIP strip) or liposome co-sedimentation assay with wild-type and FYVE-mutant CG6051 to formally confirm PI3P binding. This would provide direct experimental support for the GO:0032266 annotation.
- **Endosomal localization:** Express GFP-tagged CG6051 in *Drosophila* S2 cells or imaginal discs and confirm co-localization with early endosome markers (Rab5, Hrs/HGS).
- **EGFR trafficking assay:** Test whether CG6051 knockdown or overexpression affects EGFR degradation kinetics, consistent with its predicted negative regulatory role in EGFR signaling (IBA annotation).

Priority 4: Computational validation

- **Hidden Markov Model search:** Run CG6051 full-length sequence against Pfam phosphatase domain HMMs (PF00102/PTP, PF00782/DSPc, PF00782/myotubularin) to formally confirm absence of any cryptic phosphatase domain below standard detection thresholds.

- **AlphaFold structure analysis:** Examine the AlphaFold-predicted structure of CG6051 (AF-Q9VB70-F1) for any structural resemblance to phosphatase active sites, even in the absence of sequence-level detection.

Curation Leads

Lead 1: Remove GO:0004438 (PI3P phosphatase activity) — HIGH CONFIDENCE

- **Action:** Remove annotation
- **Rationale:** No phosphatase domain, no catalytic motif, no experimental evidence. IEA:ARBA-only annotation contradicted by ortholog data from two species.
- **Verification:** Compare domain architecture via InterPro for Q9VB70 (FYVE only, no phosphatase domains) vs. Q9VMI9/Mtm (myotubularin phosphatase domain with IDA evidence for GO:0004438). InterPro entry IPR051118 (LST-2 family) has no GO term mappings to phosphatase activity.

Lead 2: Remove GO:0052629 and GO:0004721 — HIGH CONFIDENCE

- **Action:** Remove both annotations
- **Rationale:** Same domain-architecture argument as Lead 1. All three phosphatase annotations are IEA:ARBA with no supporting non-IEA evidence.

Lead 3: Consider adding GO:0032266 (PI3P binding) — MODERATE CONFIDENCE

- **Action:** Add with ISS evidence
- **Basis:** *C. elegans* LST-2 (Q9TZD0) has GO:0032266 with ISS evidence. Human ANKFY1 (Q9P2R3) has GO:1901981 (PIP binding) with IDA evidence.
- **Candidate reference for verification:** UniProt entry for ANKFY1 (Q9P2R3) — annotated as binding PI3P based on experimental evidence.

Lead 4: Flag ARBA rule for UniProt review

- **Action:** Report to UniProt that the ARBA rule generating phosphatase annotations for FYVE-containing proteins (in the absence of a phosphatase domain) produces false annotations.
- **Rationale:** The rule may be inappropriately transferring annotations from FYVE+phosphatase domain proteins (e.g., MTMR3/MTMR4) to FYVE-only proteins (e.g., lst-2 family).

Lead 5: Suggested experiment for definitive resolution

- **Question:** Does CG6051 have any detectable phosphatase activity in vitro?
- **Experiment:** Express and purify CG6051, test against PI3P and pNPP substrates with appropriate controls (Mtm as positive, heat-inactivated CG6051 as negative).
- **Expected outcome:** No phosphatase activity, confirming the domain-based prediction.

Appendix: Annotation Provenance Summary

All phosphatase-related GO annotations on CG6051/Q9VB70 trace to a single source: **GO_REF:0000117 (UniProtKB-ARBA automatic annotation)**. No manual curation, no experimental evidence, no phylogenetic inference supports the phosphatase assignment. The two manually/phylogenetically derived annotations (IBA from GO_Central) support adaptor/endosome function, which is consistent with the FYVE domain biology and ortholog characterization. This case illustrates a known limitation of automatic annotation pipelines: feature-based rules can conflate binding domains with catalytic domains when both occur in some members of a protein superfamily but not others.
