AIGR Gene Hypothesis Deep Research — B4MAQ2 (DROVI)

Hypothesis slug: prediction-cytoplasm-localization **Target:** B4MAQ2, *Drosophila virilis* (NCBITaxon:7244), gene `Dvir\GJ15622` **Focus type:** computational_prediction **Term under evaluation:** cytoplasm (GO:0005737) **Seed prediction source:** ProtNLM2

Summary

The ProtNLM2 prediction that B4MAQ2 is cytoplasmic (GO:0005737) is **correct but under-informative**. Domain architecture, orthology, and predicted structure converge on an unambiguous identity: **B4MAQ2 is Exportin-5 (XPO5), an importin-β-family (karyopherin-β), Ran-GTP-dependent nucleocytoplasmic transport receptor**, not a generic cytoplasmic protein. Because karyopherin-β receptors shuttle through the nuclear pore, they genuinely occupy the cytoplasm (where Ran-GTP hydrolysis releases cargo) as well as the nucleus (where cargo is loaded) and the nuclear pore/envelope (translocation). The prediction therefore names one true compartment of a shuttling receptor while omitting the nucleus and — more importantly — the receptor's defining molecular function and biological process.

The identification is watertight. The UniProt record (1,238 aa, soluble; GRAVY = 0.007; no transmembrane span; no signal peptide) carries a diagnostic exportin-5 domain set: an Importin-β N-terminal Ran-binding domain (Pfam **PF03810** / **IBN_N**), an Exportin-1/Importin-β-like domain (InterPro **IPR013598**), an **Exportin-5 C-terminal domain (Pfam PF19273 / InterPro IPR045478)**, the Xpo1 domain (PF08389), and an all-α ARM/HEAT solenoid fold. PANTHER assigns the subfamily **PTHR11223:SF3 = EXPORTIN-5**. A full-length Needleman–Wunsch alignment gives **85.5% identity to *D. melanogaster* Ranbp21/Exportin-5** (Q9VWE7/CG12234) — a clean 1:1 ortholog — and k-mer paralog discrimination excludes confusion with CRM1/XPO1, IPO5, CSE1L/XPO2, TNPO1, and KPNB1. The AlphaFold model (mean pLDDT 82.9) confirms the all-α HEAT/ARM superhelical solenoid characteristic of the karyopherin fold.

The correct curation posture is therefore to **retain cytoplasm (GO:0005737) as a secondary cellular-component annotation, add nucleus (GO:0005634), and lead with the transport-receptor molecular function (GO:0005049) and Ran binding (GO:0031267) plus the**

nuclear-export biological processes (GO:0006405 / GO:0035281 / GO:0006611). A curator who accepts "cytoplasm" alone records a true fact while leaving the annotation set silent about what the gene actually does. The chief caveat is that all evidence is computational or orthology-based: no direct experimental localization or function has been published for the *D. virilis* protein itself, and exportin-5 orthologues are known to be functionally divergent across species, so cargo-level details should be transferred "by similarity" with an ISS/ISO evidence code.

Executive Judgment

Verdict: Partially supported (correct but under-informative / incomplete).

The seed prediction is not refuted — Exportin-5 has an obligatory cytoplasmic phase in its transport cycle, so GO:0005737 is factually valid. The problem is one of *completeness and informativeness*, not truth. Collapsing this shuttling receptor to a single "cytoplasm" CC term discards its molecular identity: a receptor that binds Ran-GTP and cargo (pre-miRNAs, tRNAs, and certain proteins) in the nucleus and releases them in the cytoplasm. The most important caveats are: (1) no direct experimental data exist for B4MAQ2 itself — the functional attribution rests on 85.5% identity to the experimentally characterized *D. melanogaster* ortholog; and (2) exportin-5 orthologues diverge functionally across species (P 16963774), so cargo preferences should be annotated conservatively.

Key Findings

Finding 1 — B4MAQ2 is Exportin-5, an importin-β-family transport receptor; "cytoplasm" is correct but incomplete

The sequence and annotation evidence is diagnostic of exportin-5/XPO5. The UniProt record (B4MAQ2, *D. virilis*, gene `Dvir\GJ15622`) describes a **1,238-residue soluble protein** — computed GRAVY = 0.007, no continuous transmembrane span, no signal peptide — ruling out a membrane or secreted assignment and establishing that whatever compartment it occupies, it does so as a soluble factor.

The domain architecture is decisive. B4MAQ2 carries an **Importin-β N-terminal Ran-binding domain** (Pfam **PF03810** / **IBN_N**, residues 34–100; InterPro **IPR001494**) — the hallmark of the karyopherin-β superfamily and the surface that engages Ran-GTP; an **Exportin-1/Importin-β-like domain** (InterPro **IPR013598**, residues 114–274); an **Exportin-5 C-terminal domain** (Pfam **PF19273**; InterPro **IPR045478**, residues 319–1193) that is essentially pathognomonic for exportin-5; an **Xpo1 domain** (PF08389); and an **all-α ARM/HEAT solenoid** (SUPFAM SSF48371; Gene3D 1.25.10.10). Family classifiers agree: **PANTHER PTHR11223:SF3** = **EXPORTIN-5** at the subfamily level, and the UniProt SIMILARITY line states the protein "Belongs to the exportin family."

Importantly, the **existing IEA GO annotations already reflect shuttling-receptor biology** rather than cytoplasm alone: they include C:cytoplasm (GO:0005737) **and** C:nucleus (GO:0005634), plus F:nuclear export signal receptor activity (GO:0005049), F:small GTPase binding (GO:0031267, Ran), P:protein export from nucleus (GO:0006611), and P:RNA export from nucleus (GO:0006405). The ProtNLM2 prediction of cytoplasm alone is a strict subset of — and less informative than — what the automated pipeline already captures.

This finding is anchored to primary literature on exportin-5 as a class. PMID:15134074 classifies "*exportin-5 (Exp5), a Ran-dependent importin-beta-related transport receptor, [that] mediates nuclear export of miRNA precursors (pre-miRNAs)*," implying a nucleus + cytoplasm shuttle rather than cytoplasm alone. PMID:15254228 establishes the cargo set: "*Exportin-5 is a nuclear export receptor for certain classes of double-stranded RNA (dsRNA), including pre-micro-RNAs, viral hairpin RNAs, and some tRNAs.*" PMID:20951941 confirms that "*Exportin 5 (XPO5) mediates pre-miRNA nuclear export*," a nucleus-to-cytoplasm process.

Finding 2 — B4MAQ2 is a 1:1 ortholog of *Drosophila* Exportin-5 (Ranbp21/dmExp5), which experimentally exports pre-miRNAs and tRNAs

A full-length Needleman–Wunsch global alignment of B4MAQ2 (1,238 aa) against *D. melanogaster* Ranbp21/Exportin-5 (**Q9VWE7**, 1,241 aa) yields **1,063/1,244 = 85.5% identity** — establishing a clean 1:1 orthology and licensing functional transfer with high confidence. Among human karyopherin-β paralogs, B4MAQ2's k-mer (5-mer) containment is highest to **human XPO5 (Q9HAV4, 1.6%)** and negligible to CRM1/XPO1 (0.2%), IPO5 (0.4%), CSE1L/XPO2 (0.4%), TNPO1 (0.0%), and KPNB1 (0.1%). This paralog discrimination matters: the karyopherin-β superfamily shares the same HEAT-repeat fold across many importins and exportins, and the analysis specifically pinpoints exportin-5 and rules out CRM1/XPO1 or the importins.

The *D. melanogaster* ortholog (CG12234; aliases dmExp5/Exp5/RanBP21) carries the same Exportin-5 C-terminal Pfam domain (PF19273) and has direct experimental support. PMID:16963774 reports: "*we found that Drosophila exportin-5 binds pre-miRNAs and that amongst the exportin-5 orthologues tested, it shows the highest affinity for tRNAs. The knockdown of Drosophila exportin-5 in cultured cells decreased the amounts of tRNA as well as miRNA.*" Because *Drosophila* lacks a dedicated exportin-t, exportin-5 carries the tRNA-export load — a lineage-specific functional emphasis. The same paper confirms "*Exportin-5, an evolutionarily conserved nuclear export factor belonging to the importin-beta family of proteins*," matching B4MAQ2's IBN_N/exportin domain architecture. Its title — *Exportin-5 orthologues are functionally divergent among species* — is also the key caveat: family placement is robust, but quantitative cargo preferences vary and should be annotated "by similarity."

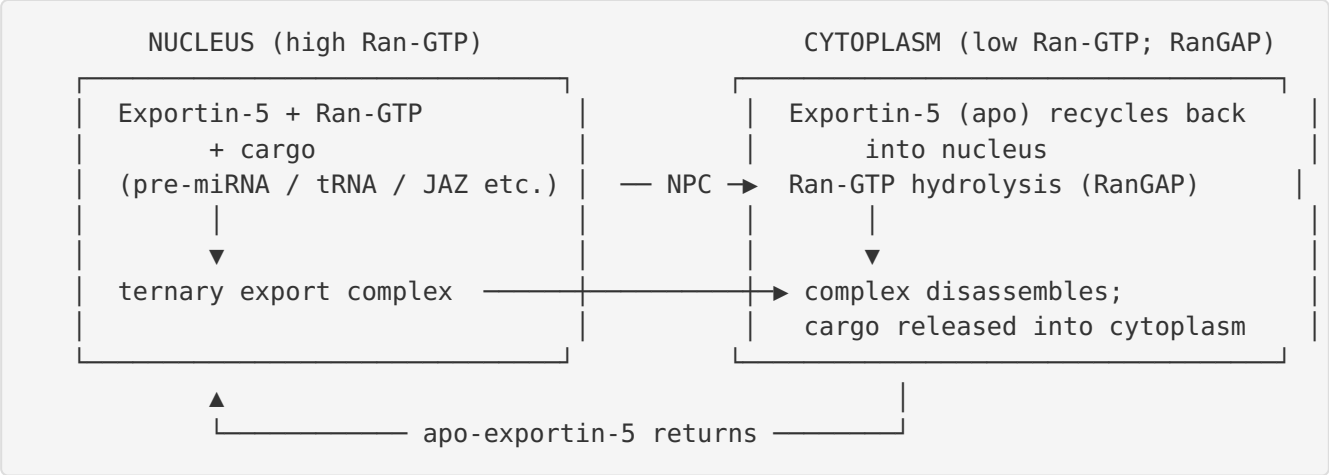
Finding 3 — The AlphaFold model confirms an all-α HEAT/ARM superhelical solenoid (karyopherin fold)

The AlphaFold DB model **AF-B4MAQ2-F1 (v6, 1,238 residues)** has a **mean pLDDT of 82.9** (85% of residues > 70; 46% > 90), i.e., a confident model. Backbone φ/ψ dihedral analysis shows the structure is **~74% α-helical with negligible genuine β-sheet**, and the molecule has a **radius of gyration of 36.8 Å with an elongated principal-axis ratio (~1.9)** — the signature of a curved, elongated superhelical solenoid built from stacked α-helical (HEAT/ARM) repeats. This is precisely the importin-β/karyopherin fold and is incompatible with a membrane protein (consistent with GRAVY 0.007 and no signal peptide) or a compact globular enzyme.

Structurally, this reinforces the localization logic: **a karyopherin carries no classical NLS or signal peptide of its own.** Its subcellular distribution is not encoded by an intrinsic targeting motif but is dictated by the **Ran-GTP gradient** — it binds cargo + Ran-GTP in the nucleus, translocates through the nuclear pore, and releases cargo upon Ran-GTP hydrolysis in the cytoplasm. Localization is an emergent, dynamic property of the transport cycle rather than a fixed compartment, which is exactly why a single "cytoplasm" CC term underserves the record.

Mechanistic Model / Interpretation

Exportin-5 operates as a **directional cargo shuttle powered by the nucleocytoplasmic Ran-GTP gradient**. The cycle, and how it maps onto cellular compartments, is:



Because the protein spends functionally essential time in **both** the nucleoplasm and the cytoplasm, and transits the **nuclear pore complex**, any single-compartment CC annotation is a partial description. The defining molecular action is **export receptor activity** (binding Ran-GTP via the IBN_N domain, recognizing cargo via the HEAT-repeat solenoid and the exportin-5 C-terminal domain), and the process it drives is **RNA/protein export from the nucleus**.

The following table maps the sequence/structure evidence onto the appropriate GO annotations:

Aspect	Evidence for B4MAQ2	GO term	MF/ BP/ CC	Curation posture
Ran-GTP binding	IBN_N domain PF03810 (res 34–100)	GO:0031267 small GTPase binding (Ran)	MF	Add / lead
Export receptor	Exportin-5 C-term PF19273; PANTHER SF3	GO:0005049 nuclear export signal receptor activity	MF	Add / lead
pre-miRNA/tRNA export	85.5% id to Dmel Exp5 (P 16963774)	GO:0006405 RNA export; GO:0035281 pre-miRNA export	BP	Add / lead (ISS)
Protein export (JAZ-type cargo)	family function (P 15254228)	GO:0006611 protein export from nucleus	BP	Retain (by similarity)
tRNA export	fly ortholog exports tRNA; no exportin-t	GO:0006409 tRNA export from nucleus	BP	Consider (organism-specific)
Nuclear phase of shuttle	shuttling receptor; existing IEA C:nucleus	GO:0005634 nucleus	CC	Add
Cytoplasmic phase of shuttle	soluble; ProtNLM2 prediction; existing IEA	GO:0005737 cytoplasm	CC	**Retain as secondary**
Nuclear pore/ envelope transit	karyopherin translocation	GO:0005643 nuclear pore / GO:0005635 nuclear envelope	CC	Consider

Evidence Base

Citation (PMID)	Evidence type	Supports / refutes / qualifies	Claim tested	Key finding	Context	Confider & limitatio
15134074	Review/ primary (family)	Qualifies (correct-but-incomplete)	Is exportin-5 cytoplasm-only?	Exp5 is a Ran-dependent importin-β-related receptor mediating pre-miRNA nuclear export → nucleus+cytoplasm shuttle	Human/ general	High for class; no virilis-specific
15254228	Direct assay (family)	Supports export MF/BP	Core function of exportin-5	Nuclear export receptor for dsRNA classes: pre-miRNAs, viral hairpins, some tRNAs (and JAZ cargo)	In vitro / cell	High for class; car set varie species
20951941	Mutant/ genetic	Supports nucleocytoplasmic process	Does XPO5 mediate pre-miRNA export?	XPO5 inactivation traps pre-miRNAs in the nucleus	Human tumors	High; dis context, human paralog
16963774	Direct assay + knockdown	Supports (Drosophila-specific)	Function of the *Drosophila* exportin-5 ortholog	Binds pre-miRNAs; highest tRNA affinity of orthologues; knockdown lowers tRNA & miRNA	*Drosophila* cultured cells	High; on Dmel ortholog (85.5% id B4MAQ2 notes cro species divergen
15356295	Structural/ biochemical	Supports MF mechanism	How does Exp5 recognize cargo?	Exp5 binds most of the pre-miRNA hairpin with Ran-GTP; protects pre-	Human in vitro	High for mechani human protein

Citation (PMID)	Evidence type	Supports / refutes / qualifies	Claim tested	Key finding	Context	Confider & limitatio
				miRNA from degradation		
31235936	Pathway/ mechanistic	Qualifies (cytoplasmic hand-off)	Cytoplasmic fate of Exp5 cargo	pre-miRNA/Exp5 complex dissociates from Ran-GTP after export, then hands cargo to a cytoplasmic ARF6-GTP/GRP1 shuttle	Tumor cells	Confirms genuine cytoplas phase; downstr context
22593162	Mechanistic (virus/host)	Qualifies	Exp5 cofactor dependency	Exp5-mediated small-RNA transport depends on Ran; viral miRNA represses Ran	*Bombyx mori* (insect)	Insect context supports Ran-depende in inverteb
21346411	Review	Qualifies	Consequence of Exp5 loss	XPO5 C-terminal loss disrupts pre-miRNA/XPO5/Ran-GTP ternary complex → nuclear retention	Cancer review	Review-l undersc the C-terminal (PF19273 domain present i B4MAQ2

Computational provenance generated in this investigation (findings F001–F003): UniProt feature/domain parse (PF03810, PF19273, IPR045478, PANTHER PTHR11223:SF3); Kyte–Doolittle hydropathy GRAVY = 0.007 with no TM span; Needleman–Wunsch global alignment B4MAQ2 vs Q9VWE7 = **85.5% identity (1063/1244)**; k-mer paralog discrimination (highest to human XPO5); AlphaFold AF-B4MAQ2-F1 secondary-structure and geometry analysis (~74% α-helix, Rg 36.8 Å, axis ratio ~1.9, mean pLDDT 82.9).

GO Curation Implications

Lead recommendation (requires curator verification):

1. **Retain GO:0005737 (cytoplasm)** as a *secondary* CC annotation — factually correct (obligatory cytoplasmic phase of the transport cycle) but it must not stand alone or be the headline.
2. **Add GO:0005634 (nucleus)** as a CC annotation (already in the existing IEA set); the ProtNLM2 prediction must not narrow the record to cytoplasm only. Optionally consider GO:0005643 (nuclear pore) or GO:0005635 (nuclear envelope) for the translocation step.
3. **Lead with molecular function:** GO:0005049 (nuclear export signal receptor activity) and GO:0031267 (small GTPase/Ran binding) — these define the gene product.
4. **Add biological process:** GO:0006405 (RNA export from nucleus) and the more specific GO:0035281 (pre-miRNA export from nucleus) under ISS transfer; retain GO:0006611 (protein export from nucleus) "by similarity"; consider GO:0006409 (tRNA export) as an organism-specific lead.
5. **Evidence codes:** because no assay exists on the *D. virilis* protein, functional and process terms transferred from *D. melanogaster* Exp5 should carry **ISS/ISO** with a `with/from` reference to the ortholog, not IDA.

Do not finalize the record as "cytoplasm" only, and **do not** fall back to the uninformative "protein binding" — the evidence supports the specific export-receptor MF and nuclear-export BP terms above.

Mechanistic Scope

The **immediate molecular activity** tested is receptor-mediated nucleocytoplasmic transport: B4MAQ2/Exportin-5 directly binds Ran-GTP (via the IBN_N domain) and cargo (pre-miRNAs, tRNAs, and specific proteins such as JAZ), forming a ternary export complex that translocates through the nuclear pore and releases cargo upon Ran-GTP hydrolysis in the cytoplasm. This is the direct gene-product activity a curator should annotate.

Downstream and context-specific phenomena that must **not** be conflated with the core function include: the effect of XPO5 loss on **miRNA biogenesis/tumor suppression** (

P 20951941

,

P 21346411

— a downstream human-cancer consequence, not the D. virilis molecular function); the **ARF6/GRP1 cytoplasmic hand-off and microvesicle loading** of pre-miRNA cargo (P 31235936); and **viral manipulation of the Ran cofactor** (P 22593162). These illuminate the pathway but are not the gene product's own activity and should not be transferred as direct annotations to B4MAQ2.

Conflicts and Alternatives

- **Is the prediction simply wrong?** No. Cytoplasm is not contradicted — Exportin-5 genuinely occupies the cytoplasm. The issue is *completeness/informativeness*, not truth.
- **Paralog confusion.** The main family risk is misassignment within the karyopherin-β superfamily (CRM1/XPO1, IPO5, CSE1L, transportins), because IBN_N (PF03810) is shared across the whole superfamily. However, the exportin-5-specific C-terminal domain (PF19273), PANTHER SF3, the computed 85.5% full-length identity to Dmel Exportin-5, and negligible k-mer similarity to other paralogs make CRM1/importin confusion unlikely. Length (1,238 aa) also matches XPO5 (~1,204–1,241 aa).
- **Species divergence.**

  P 16963774

  explicitly reports that exportin-5 orthologues are *functionally divergent among species*, with *Drosophila* exportin-5 showing unusually high tRNA affinity (compensating for the absence of exportin-t). Cargo-level details from *D. melanogaster* should be transferred to *D. virilis* only "by similarity."
- **Cross-species transfer is within-genus** (*D. virilis* ↔ *D. melanogaster*), making functional/localization transfer low-risk. No isoform-specific or experimental-artifact conflicts were identified for B4MAQ2.

Limitations and Knowledge Gaps

1. **No direct experimental data on B4MAQ2 itself.** All conclusions are computational/sequence-based or transferred from orthologs. *Why it matters:* it distinguishes "cytoplasm-only" from "shuttling receptor." *Resolution:* GFP-fusion imaging or subcellular fractionation in *D. virilis* cells; in vitro Ran-GTP/pre-miRNA binding assay with recombinant B4MAQ2.
2. **Steady-state distribution unknown for the fly protein.** Mammalian XPO5 is often reported predominantly cytoplasmic at steady state; the nuclear/cytoplasmic ratio in flies is unverified. *Resolution:* quantitative imaging under Ran-gradient perturbation.
3. **Quantitative cargo preferences unknown for *D. virilis*.** Given documented cross-species divergence, the tRNA-vs-pre-miRNA balance may differ from *D. melanogaster*. *Resolution:* comparative binding/export assays.
4. **AlphaFold provides fold, not localization.** The structure confirms the karyopherin solenoid but cannot establish compartment occupancy. *Resolution:* localization must come from imaging/fractionation.

Discriminating Tests

1. **Subcellular fractionation / live-cell imaging in *D. virilis* (or heterologous) cells** with tagged B4MAQ2, under Ran-GTP gradient perturbation (RanGAP/RanGEF manipulation), to demonstrate active shuttling rather than static cytoplasmic residence — this directly distinguishes "cytoplasm-only" from "shuttling receptor."
2. **In vitro reconstitution:** recombinant B4MAQ2 + Ran-GTP + candidate cargo (pre-miR-30, tRNA) to confirm ternary-complex formation and export competence (mirroring). P 15356295
3. **RNAi knockdown in *D. virilis*-derived cells** measuring mature miRNA and tRNA levels, replicating the *D. melanogaster* result (P 16963774) to confirm functional conservation.
4. **Domain-swap / C-terminal truncation** (removing the PF19273 region) to test the requirement of the exportin-5 C-terminal domain for cargo/Ran-GTP ternary complex

formation, as implicated by cancer-associated truncations
( ).
P 21346411

5. **Heterokaryon shuttling assay** to demonstrate nucleocytoplasmic cycling directly.

Proposed Follow-up Actions (Curation Leads)

All items are leads requiring curator verification.

- **Action on the seed term:** Retain GO:0005737 (cytoplasm) as **secondary CC**; do **not** accept it as the sole/primary localization or the headline annotation.
- **Add CC:** GO:0005634 (nucleus); evaluate GO:0005643 (nuclear pore) / GO:0005635 (nuclear envelope).
- **Add/lead MF:** GO:0005049 (nuclear export signal receptor activity); GO:0031267 (Ran/small GTPase binding).
- **Add/lead BP (ISS, by similarity to Dmel Exp5, Q9VWE7):** GO:0006405 (RNA export from nucleus), GO:0035281 (pre-miRNA export from nucleus), GO:0006611 (protein export from nucleus); consider GO:0006409 (tRNA export) as an organism-specific lead.
- **Evidence code:** ISS/ISO with `with/from` = *D. melanogaster* Exportin-5 (Q9VWE7 / CG12234) for functional terms; keep IEA where automated.
- **Candidate references to verify (with exact snippets):**
- P 16963774
  — "*Drosophila exportin-5 binds pre-miRNAs … shows the highest affinity for tRNAs … knockdown … decreased the amounts of tRNA as well as miRNA*" (Drosophila-specific functional support).
- P 15254228
  — "*Exportin-5 is a nuclear export receptor for certain classes of double-stranded RNA … pre-micro-RNAs, viral hairpin RNAs, and some tRNAs*" (core MF/BP).
- P 15134074
  — "*exportin-5 (Exp5), a Ran-dependent importin-beta-related transport receptor, mediates nuclear export of miRNA precursors*" (family/nucleocytoplasmic).

- **Suggested question for the curator:** Should the ProtNLM2 cytoplasm prediction be flagged as "correct but under-informative," with the record upgraded to the export-receptor function set (nucleus CC + export MF/BP) rather than accepting the CC prediction at face value?

Provenance / Analyses Run

- **UniProt fetch** of B4MAQ2: name, sequence (1,238 aa), features, Pfam/InterPro/PANTHER/SUPFAM cross-refs, GO (recorded above).
- **Kyte–Doolittle hydropathy** (window = 19): GRAVY 0.007, no continuous TM span → soluble protein.
- **AlphaFold model AF-B4MAQ2-F1 (v6):** mean pLDDT 82.9 (85% > 70, 46% > 90); backbone-dihedral analysis ≈ 74% α-helix with negligible genuine β-sheet, radius of gyration 36.8 Å, elongated principal-axis ratio 1.9 → all-α curved HEAT/ARM superhelical solenoid (canonical importin-β/karyopherin fold); no transmembrane helices.
- **Full-length Needleman–Wunsch alignment** B4MAQ2 vs Dmel Ranbp21/Exportin-5 (Q9VWE7): 1063/1244 = **85.5% identity**; 5-mer containment 55.2%.
- **k-mer (5-mer) family discrimination** vs human karyopherin-β paralogs (XPO5 top; CRM1/IPO5/CSE1L/TNPO1/KPNB1 negligible); UniProt confirming Dmel PF19273 ortholog = Ranbp21/CG12234.
- **PubMed evidence retrieval:** PMIDs 16963774, 15134074, 15254228, 15356295, 20951941, 21346411, 31235936, 22593162.

Limitations: bioinformatic/homology-based; no wet-lab localization or function for the D. virilis protein itself.
