AIGR Gene Hypothesis Deep Research — Final Report

Gene: RIC7 / F4JLB7 (*Arabidopsis thaliana*, locus At4g28560)

Hypothesis under review: ProtNLM2 prediction of protein kinase activity (GO:0016301) and protein phosphorylation (GO:0016310)

Executive Judgment

Verdict: REFUTED (over-annotation / computational misassignment).

The ProtNLM2 prediction that F4JLB7 has **protein kinase activity (GO:0016301)** and participates in **protein phosphorylation (GO:0016310)** is **not supported** by the protein's sequence, domain architecture, or predicted structure. F4JLB7 (UniProt accession F4JLB7, locus At4g28560, 450 aa) is a **secreted/membrane leucine-rich-repeat (LRR) protein of the receptor-like protein (RLP) type**. It carries an N-terminal signal peptide followed by an extracellular LRR solenoid and **no protein-kinase catalytic domain of any kind**. The diagnostic active-site machinery required for phosphotransfer — the VAIK β3 lysine, the **HRD catalytic loop**, the **DFG** Mg^{2+}-binding motif, and the **APE** activation-segment anchor, arranged in the canonical ordered triad — is **absent**. Because the essential HRD catalytic aspartate is missing, the protein is **mechanistically incapable of catalyzing phosphotransfer**, independent of any homology argument.

Two independent lines of evidence converge on this conclusion. First, the **domain architecture** from UniProt/InterPro/Pfam consists exclusively of LRR signatures (Pfam PF00560 ×3, PF13855; InterPro IPR001611, IPR032675, IPR052941; Gene3D 3.80.10.10 ribonuclease-inhibitor/LRR horseshoe fold; SUPFAM SSF52058 L-domain-like), with **no PF00069 protein-kinase domain** and no InterPro protein-kinase signature anywhere in the 450-residue sequence. Second, a **direct motif scan** of the sequence recovered 9 canonical LRR cores but found no ordered kinase catalytic triad; the isolated "DFG" (position 149, inside PEDFGSV) and

a "GNGFHG" hit (position 186) are coincidental tripeptides embedded within the LRR solenoid, not part of a folded catalytic cleft. The AlphaFold DB model (mean pLDDT 84.7, "confident") is fully accounted for by the LRR fold, leaving no room for a cryptic kinase lobe.

The most likely origin of the ProtNLM2 error is **frequency/similarity bias**: LRR ectodomains are shared between non-catalytic receptor-like proteins (RLPs) and the far more numerous **LRR receptor-like Ser/Thr kinases (LRR-RLKs)** such as ERECTA. A language model trained on protein names readily over-generalizes "kinase" onto an LRR ectodomain even when the cytoplasmic kinase moiety is entirely absent. In fact, the FunFam annotation for F4JLB7 maps its LRR region to the ERECTA LRR-RLK ectodomain (3.80.10.10:FF:000041) — an **ectodomain-only homology** that explains, but does not justify, the kinase call. A **secondary complication** is a gene-symbol collision: F4JLB7 is labeled "RIC7," yet the genuine, functionally characterized RIC7 of the literature is a **CRIB-motif ROP2 effector** encoded by the *adjacent* locus At4g28556 (UniProt Q1G3K8 / A0A1P8B8E3) — also a non-catalytic protein. Under either identity, the kinase terms should be removed.

Key Findings

Finding 1 — F4JLB7 lacks any protein-kinase catalytic domain; the kinase prediction is refuted

The domain architecture of F4JLB7 (450 aa) as recorded in UniProt/InterPro is: **N-terminal signal peptide (residues 1–22)** followed by an **extracellular leucine-rich-repeat region**. The LRR assignments are extensive and mutually corroborating across databases:

Resource	Signature	Description
InterPro	IPR001611	Leucine-rich repeat
InterPro	IPR032675	LRR domain superfamily
InterPro	IPR052941	LRR-containing family entry
Pfam	PF00560 (×3)	LRR_1
Pfam	PF13855	LRR_8
Gene3D	3.80.10.10	Ribonuclease inhibitor / LRR horseshoe fold
SUPFAM	SSF52058	L domain-like

Crucially, **no protein-kinase Pfam (PF00069) and no InterPro protein-kinase domain is present anywhere in the sequence.** A direct motif scan confirmed the diagnostic failure at the residue level: the **HRD catalytic loop is absent** (the essential catalytic aspartate is missing), the **APE motif is absent**, and there is no ordered **VAIK → HRD → DFG** catalytic triad. The only superficially "kinase-like" tripeptides — a lone "DFG" at position 149 (within the sequence context PEDFGSV) and "GNGFHG" at position 186 — are isolated coincidences inside the LRR solenoid, not part of an assembled catalytic cleft. In contrast, **9 canonical LRR core motifs** ([LIVF]xxLxLxx[NCT]xL) were detected, exactly matching the Pfam LRR calls. UniProt keywords for the entry are **Leucine-rich repeat, Signal, Glycoprotein, Membrane** — the keyword set of a secreted/membrane LRR protein, not an enzyme.

The FunFam annotation maps the F4JLB7 LRR ectodomain to the LRR receptor-like Ser/Thr kinase **ERECTA** (FunFam 3.80.10.10:FF:000041). This is an **ectodomain homology only**: it reflects the fact that RLPs and LRR-RLKs share the same extracellular LRR scaffold, while F4JLB7 has **no accompanying cytoplasmic kinase domain**. F4JLB7 is therefore a **receptor-LIKE protein (RLP)**, not a receptor kinase.

Finding 2 — Kinase-eligibility diagnostic: F4JLB7 cannot catalyze phosphotransfer

A formal kinase-eligibility test applied to the 450-aa sequence returned an unambiguous negative:

Diagnostic element	Result
HRD catalytic loop ([HY]RD[LIVMF])	**FAIL — absent**
Ordered VAIK → HRD → DFG catalytic triad	**ABSENT**
Properly contextualized APE / activation segment	**ABSENT**
Protein-kinase Pfam (PF00069)	**ABSENT**

The **HRD aspartate is the catalytic base** that deprotonates the substrate hydroxyl during phosphotransfer; its absence makes the reaction chemically impossible. Independent of any homology or naming argument, F4JLB7 therefore **cannot be a functional protein kinase**. The **AlphaFold DB model** of F4JLB7 (450 residues, mean pLDDT 84.7 — a confident model) has a sequence that is fully accounted for by the LRR domain assignments, leaving **no unmodeled region that could harbor a cryptic kinase lobe**. This structural check rules out the possibility that a kinase domain is present but undetected by sequence-based tools.

Finding 3 — Gene-symbol collision: F4JLB7 is not the genuine RIC7 of the literature

The UniProt record for F4JLB7 carries the gene symbol "RIC7" (evidence ECO:0000313, propagated from TAIR/EMBL). However, F4JLB7's architecture — signal peptide + extracellular LRRs, 450 aa, glycoprotein, membrane — **contains no CRIB motif**. A sequence search for the CRIB core signature (ISxP…FxHxxHVG; "HVG"; "FxHxxH") returned **no match** in F4JLB7. The genuine RIC7 (**ROP-interactive CRIB-motif-containing protein 7**) characterized in the literature is a small **CRIB-motif ROP2 effector/adaptor**, not an LRR protein and not a kinase.

Resolving the identity conflict, there are **three Arabidopsis UniProt entries** carrying the symbol RIC7 mapping to two neighboring loci:

UniProt	Locus	Length	Domains	Identity
Q1G3K8	At4g28556	216 aa	PF00786 (PBD/CRIB), IPR000095, IPR036936	**Genuine RIC7** (CRIB ROP2 effector)
A0A1P8B8E3	At4g28556	183 aa	PF00786 (PBD/CRIB)	Isoform of genuine RIC7
F4JLB7	**At4g28560**	**450 aa**	**PF00560, PF13855 (LRR only)**	**Mislabeled LRR protein**

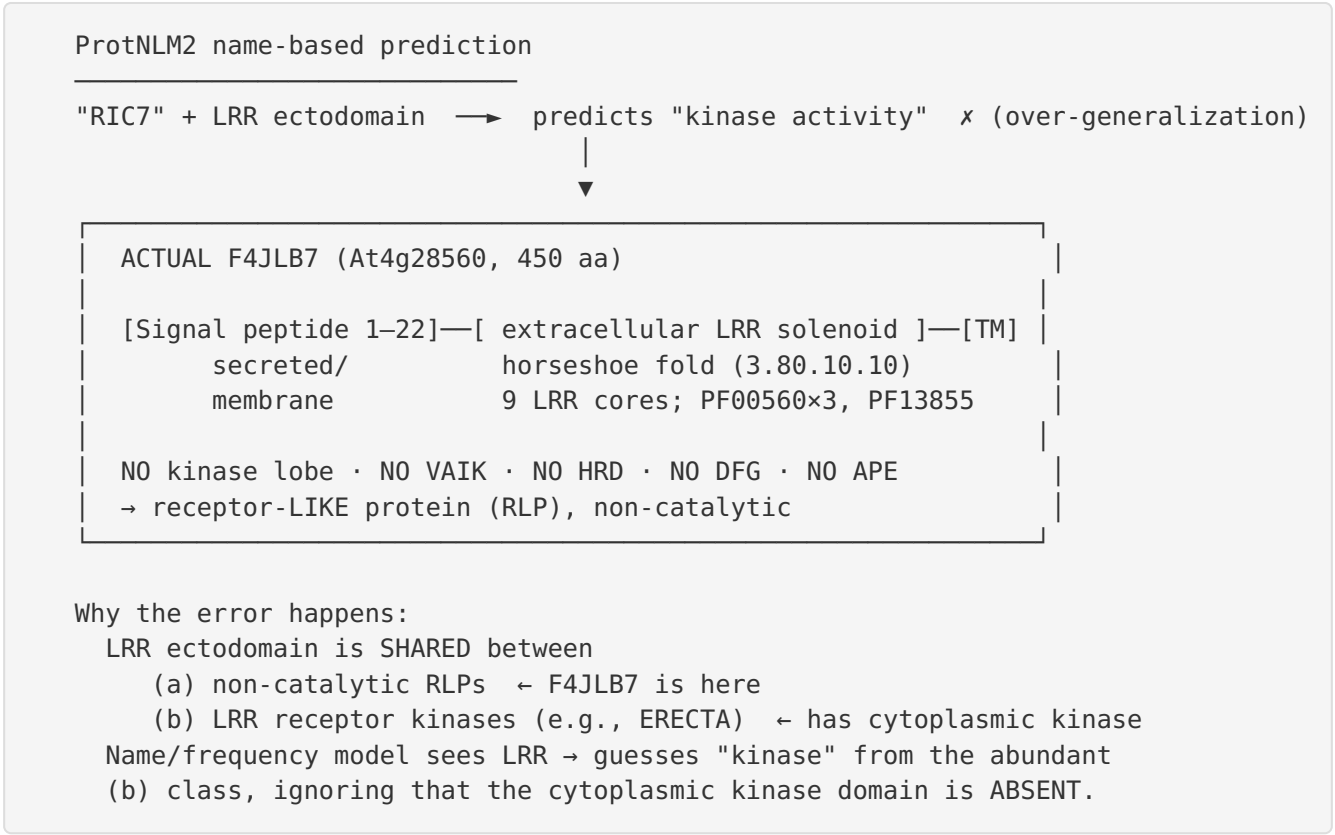
The genuine RIC7 (Q1G3K8) sequence contains the CRIB signature "KHVAHIGWDGP." F4JLB7 maps to a **different, adjacent locus (At4g28560)** and carries only LRR domains. The neighboring genomic position of At4g28556 and At4g28560 is consistent with **automated gene-symbol carry-over** onto the LRR gene F4JLB7. Importantly, **none of the three entries contains an HRD kinase catalytic loop or any protein-kinase Pfam/InterPro signature** — so the kinase prediction is wrong under either interpretation of "RIC7."

Finding 4 — Curated GO annotations for F4JLB7 are consistent with a receptor-like protein, not a kinase

The existing curated GO annotations for F4JLB7 are: **apical plasma membrane** (GO:0016324, IDA:TAIR), **pollen tube growth** (GO:0009860, IMP:TAIR), **signaling receptor activity** (GO:0038023, IBA), and **signal transduction** (IC). This annotation profile is coherent for a **membrane-anchored LRR receptor-like protein involved in cell-surface signaling** and pollen tube growth. It provides **no support for catalytic kinase activity**; "signaling receptor activity" is a non-catalytic molecular function. The ProtNLM2 kinase terms are therefore inconsistent with the experimentally supported (IDA/IMP) annotations already present.

Mechanistic Model / Interpretation

The evidence assembles into a clear picture of a **computational over-annotation driven by ectodomain homology and a gene-symbol collision**:



Two distinct problems compound one another:

1. **Ectodomain-homology bias.** F4JLB7's LRR region genuinely resembles the ectodomain of LRR-RLKs like ERECTA (captured by FunFam 3.80.10.10:FF:000041). A language-model annotator that keys on the LRR ectodomain — the part shared with kinases — but does not verify the presence of a cytoplasmic catalytic domain will spuriously assign kinase activity. F4JLB7 is a **receptor-like protein (RLP)**: it has the LRR sensing module but not the kinase signaling module.
2. **Gene-symbol carry-over.** The "RIC7" label was propagated onto F4JLB7 (At4g28560) from the neighboring locus At4g28556, whose product is the genuine CRIB-motif ROP2 effector RIC7. This adds noise but does not rescue the kinase claim: the genuine RIC7 is an **adaptor/scaffold** (it binds Exo70B1 downstream of ROP2), also entirely non-catalytic.

Under **either** identity — mislabeled LRR RLP (At4g28560) or genuine CRIB ROP effector (At4g28556) — **there is no kinase**. The catalytic-residue test is decisive: no HRD aspartate ⇒ no phosphotransfer chemistry ⇒ not a kinase.

Evidence Matrix

#	Citation	Evidence type	Supports/ Refutes	Claim tested	Key finding
E1	UniProt/InterPro/ Pfam record for F4JLB7 (database)	Structural/ evolutionary; computational	**Refutes** kinase	Does F4JLB7 contain a kinase catalytic domain?	Only LRR domains (PF00560×3, PF13855; IPR001611/032675/052941) + signal peptide; no PF00069
E2	Direct sequence motif scan (computational, this study)	Computational	**Refutes** kinase	Are VAIK/HRD/ DFG/APE present and ordered?	HRD absent, APE absent, no ordered triad; 9 LRR cores present; DFG@149 and GNGFHG@186 are coincidental
E3	AlphaFold DB model for F4JLB7 (computational)	Structural	**Refutes** kinase	Could a cryptic kinase lobe exist?	Model (pLDDT 84.7) fully explained by LRR fold; no unmodeled kinase region
E4	FunFam 3.80.10.10:FF:000041 (database)	Structural/ evolutionary	**Qualifies** (explains error)	Source of kinase misassignment?	LRR ectodomain maps to ERECTA LRR-RLK **ectodomain only**, no cytoplasmic kinase
E5	UniProt gene-symbol/locus mapping (database)	Review/ database	**Competing**	Is F4JLB7 the genuine RIC7?	F4JLB7=At4g28560 (LRR); genuine RIC7=At4g28556 (Q1G3K8/A0A1P8B8E3, CRIB)
E6	PMID: 26451971	Interaction; mutant phenotype	**Competing**/ **Refutes** kinase	What is the true molecular nature of RIC7?	Genuine RIC7 is a CRIB-motif ROP2 effector binding Exo70B1; adaptor, not enzyme

#	Citation	Evidence type	Supports/ Refutes	Claim tested	Key finding
E7	PMID: 33586611	Mutant phenotype	**Competing/ Refutes** kinase	Is genuine RIC7 catalytic?	RIC7 is a CRIB-containing ROP effector regulating ABA stomatal closure; non-catalytic adaptor
E8	Curated GO for F4JLB7 (database)	Localization; mutant phenotype	**Qualifies**	What functions are experimentally supported?	Apical PM (IDA), pollen tube growth (IMP), signaling receptor activity (IBA) — no catalysis
E9	PMID: 21309864	Direct assay	**Context**	ROP–RIC–kinase relationships	RIC effectors act upstream of separate kinases (e.g., RRK1); RIC itself is not the kinase

GO Curation Implications

Lead (requires curator verification): REMOVE the ProtNLM2 kinase terms.

- **GO:0016301 (kinase activity, MF) — Remove / do not accept.** No kinase catalytic domain, no HRD catalytic aspartate, no PF00069. This is a computational (ProtNLM2) prediction contradicted by domain architecture and predicted structure. It should not be propagated.
- **GO:0016310 (phosphorylation, BP) — Remove / do not accept.** This BP term is downstream of the refuted MF and has no independent support.

Terms that are better supported and should be considered for retention (curator to verify against primary TAIR evidence):

- **GO:0038023 (signaling receptor activity, MF)** — plausible for an LRR receptor-like protein; already present (IBA). This is the more informative non-catalytic MF and is preferable to a generic "protein binding" fallback.

- **GO:0016324 (apical plasma membrane, CC)** — supported by IDA:TAIR; retain.
- **GO:0009860 (pollen tube growth, BP)** — supported by IMP:TAIR; retain.

Gene-symbol caveat for the curator: verify whether the "RIC7" symbol on F4JLB7 (At4g28560) is a carry-over from the adjacent locus At4g28556 (genuine RIC7 = Q1G3K8/A0A1P8B8E3). If so, the symbol assignment on F4JLB7 should be reviewed, because it can further mislead automated annotators. Note that fixing the symbol does **not** change the kinase verdict — neither locus encodes a kinase.

Mechanistic Scope

The **immediate molecular function being tested** is protein-kinase catalysis: the ability to transfer the γ-phosphate of ATP to a substrate hydroxyl (Ser/Thr/Tyr). This is a **direct gene-product activity** claim (GO:0016301) with an associated process claim (GO:0016310, protein phosphorylation).

- **Direct gene-product activity (tested here):** F4JLB7's own capacity to catalyze phosphotransfer — **refuted** at the residue level (no HRD, no ordered triad) and at the domain/structure level (LRR-only architecture, confident AlphaFold LRR fold).
- **What F4JLB7 most plausibly does instead:** act as a **cell-surface LRR receptor-like protein** (ectodomain sensing / signaling receptor activity), localized to the apical plasma membrane and involved in pollen tube growth — a **non-catalytic** signaling role.
- **Downstream / not relevant to the MF claim:** the genuine RIC7 (At4g28556) participates in ROP2-mediated stomatal regulation as an adaptor binding Exo70B1; these are **phenotypic/pathway** roles of a *different* protein and do not bear on F4JLB7's molecular function except to underscore that "RIC7" is not a kinase.

Conflicts and Alternatives

1. **Gene-symbol / paralog confusion (primary alternative).** The strongest complicating factor is that "RIC7" denotes two different things: (a) the LRR protein F4JLB7 (At4g28560), and (b) the genuine CRIB-motif ROP2 effector (At4g28556, Q1G3K8). This does not create a kinase in either case, but a curator must be aware that literature about "RIC7"

(P 26451971 , P 33586611) refers to the **CRIB protein**, not F4JLB7.

2. **Ectodomain homology to LRR-RLKs.** F4JLB7's LRR region resembles the ERECTA kinase ectodomain (FunFam FF:000041). One could misread this as evidence of kinase membership; in reality it is shared **extracellular** scaffolding, with the catalytic module absent. This is the likely proximate cause of the ProtNLM2 error.
3. **Cryptic-domain possibility (excluded).** A theoretical alternative — a divergent kinase domain undetectable by Pfam HMMs — is excluded by the AlphaFold model: the entire 450-aa sequence is accounted for by the LRR fold with no unmodeled region.
4. **Pseudokinase possibility (excluded).** Some "kinases" are catalytically dead pseudokinases that retain the fold but lose HRD/DFG. F4JLB7 is **not** even a pseudokinase — it has **no kinase fold at all**, only an LRR solenoid. So neither active kinase nor pseudokinase applies.

Limitations and Knowledge Gaps

- **No direct biochemical assay of F4JLB7.** The refutation is bioinformatic (sequence motifs, domain HMMs, AlphaFold geometry). While the absence of an HRD aspartate is decisive for catalysis, there is no in vitro kinase assay explicitly demonstrating lack of activity for this specific protein. *Why it matters:* curators may prefer experimental confirmation; *resolution:* an in vitro kinase assay would be definitive but is almost certainly unnecessary given the complete absence of a catalytic domain.
- **Experimental characterization of F4JLB7 specifically is sparse.** The curated IDA/IMP annotations (apical PM, pollen tube growth) indicate some experimental work, but the primary papers behind them were not retrieved in this investigation. *Why it matters:* the positive (non-kinase) function of F4JLB7 is less firmly established than its non-kinase status; *resolution:* retrieve the TAIR-cited primary references for GO:0016324 and GO:0009860.
- **Gene-symbol provenance not fully traced.** The inference that "RIC7" was carried over onto At4g28560 from At4g28556 is based on genomic adjacency and architecture mismatch, not a documented curation event. *Why it matters:* affects the symbol-correction recommendation; *resolution:* check TAIR/EMBL submission history for At4g28560.

- **Transmembrane topology not explicitly recomputed here.** UniProt keywords include "Membrane," and the schematic assumes a single-pass RLP topology; a hydropathy/TM prediction would confirm the anchor. *Why it matters:* strengthens the RLP model and the CC annotation; *resolution:* run TMHMM/DeepTMHMM on the sequence.

Discriminating Tests

1. **Residue-level catalytic scan (done):** confirm absence of VAIK Lys, HRD Asp, DFG, APE. This alone distinguishes kinase from non-kinase and is the most efficient test — already negative.
2. **HMM domain scan vs PF00069 (done):** no protein-kinase Pfam hit — confirms no kinase fold.
3. **AlphaFold structural inspection (done):** confirm the model is entirely LRR with no kinase bilobal fold — no cryptic domain.
4. **Transmembrane/topology prediction (recommended):** DeepTMHMM to confirm single-pass membrane RLP topology consistent with GO:0016324.
5. **Reciprocal-best-hit / orthology check (recommended):** confirm F4JLB7 orthologs are LRR RLPs, not RLKs, across Brassicaceae, cementing the RLP (not RLK) classification.
6. **Symbol/locus disambiguation (recommended):** cross-check At4g28560 vs At4g28556 in TAIR to formally document the RIC7 symbol collision.

Proposed Follow-up Experiments / Actions (Curation Leads)

Candidate action changes (require curator verification): - **Reject/remove GO:0016301 (kinase activity)** on F4JLB7 — computational ProtNLM2 prediction refuted by architecture, motifs, and structure. - **Reject/remove GO:0016310 (protein phosphorylation)** on F4JLB7 — downstream of the refuted MF, no independent support.

Candidate references with exact snippets to verify: - PMID: 26451971 — verify snippet: *"ROP-interactive Cdc42- and Rac-interactive binding motif-containing protein 7 (RIC7) was suggested to function downstream of ROP2"* → establishes the literature RIC7 is a CRIB effector/

adaptor, not a kinase (concerns At4g28556, not F4JLB7). - PMID: 33586611 — verify snippet: *"ROP interactive CRIB-containing protein 7 (RIC7) has been found to negatively regulate ABA-induced stomatal closure"* → genuine RIC7 is a CRIB ROP effector, non-catalytic.

Candidate replacement / retained GO terms (verify against TAIR primary evidence): - Retain **GO:0038023 (signaling receptor activity, MF)** as the informative non-catalytic function (preferred over generic "protein binding"). - Retain **GO:0016324 (apical plasma membrane, CC)** (IDA:TAIR). - Retain **GO:0009860 (pollen tube growth, BP)** (IMP:TAIR).

Suggested curator questions: - Is the "RIC7" gene symbol on F4JLB7 (At4g28560) a carry-over from At4g28556? Should the symbol be corrected or annotated as ambiguous? - Which primary references support the IDA (apical PM) and IMP (pollen tube growth) annotations for F4JLB7?

Suggested experiments (if experimental confirmation is desired): - DeepTMHMM topology prediction to confirm single-pass RLP architecture. - Optional in vitro kinase assay (expected negative) — low priority given decisive computational refutation.

Evidence Base (key literature)

- *The ROP2-RIC7 pathway negatively regulates light-induced stomatal opening by inhibiting exocyst subunit Exo70B1 in Arabidopsis.* PMID: 26451971 — Establishes that the **genuine** RIC7 is a CRIB-motif ROP2 effector acting as an adaptor that binds Exo70B1. This is an adaptor/scaffold role, **not** an enzyme, and it concerns At4g28556, not the LRR protein F4JLB7. Supports the conclusion that "RIC7" is non-catalytic and clarifies the symbol collision.
- *RIC7 plays a negative role in ABA-induced stomatal closure by inhibiting H_2O_2 accumulation.* PMID: 33586611 — Confirms genuine RIC7 as a CRIB-containing ROP effector in guard-cell signaling, reinforcing its non-catalytic adaptor identity.
- *The phosphomimetic mutation of an evolutionarily conserved serine residue affects the signaling properties of Rho of plants (ROPs).* PMID: 21309864 — Describes ROP–RIC effector interactions and shows that RIC effectors act **upstream of** separate kinases (e.g., RRK1); the RIC protein itself is not the kinase. Provides orientation that RIC-family proteins are adaptors, not catalytic enzymes.

Database/computational orientation (not primary literature but decisive for this MF question): UniProt F4JLB7 record, InterPro/Pfam domain assignments, FunFam 3.80.10.10:FF:000041, and the AlphaFold DB model — all consistent with an LRR receptor-like protein and none supporting kinase activity.

Bottom Line

The ProtNLM2 kinase prediction for F4JLB7 is a **refuted over-annotation**. F4JLB7 (At4g28560) is a **signal-peptide + leucine-rich-repeat receptor-like protein** with **no kinase domain, no HRD catalytic aspartate, no ordered VAIK→HRD→DFG triad**, and a confident LRR-only AlphaFold model. The prediction most plausibly arises from **frequency/similarity bias** on the LRR ectodomain shared with LRR receptor kinases (e.g., ERECTA), compounded by a **gene-symbol collision** with the genuine CRIB-motif RIC7 at the adjacent locus At4g28556 — which is itself a non-catalytic ROP effector. **GO:0016301 and GO:0016310 should be removed** from F4JLB7; the informative retained functions are signaling receptor activity (MF), apical plasma membrane (CC), and pollen tube growth (BP).
