AIGR Gene Hypothesis Deep Research — Final Report

Gene: *Arabidopsis thaliana* CRY1 (Cryptochrome 1, UniProt Q43125)

Hypothesis under evaluation: CRY1 has deoxyribodipyrimidine photo-lyase activity (GO:0003904)

Focus type: function_assignment · Annotation context: IBA / GO_REF:0000033 · Source: PANTHER PTHR11455

Summary

Verdict: REFUTED (family over-annotation). The seed hypothesis — that Arabidopsis CRY1 directly possesses deoxyribodipyrimidine photo-lyase activity (GO:0003904) — is not supported by primary experimental data and is directly contradicted by two independent biochemical assays on purified protein. The annotation is an inferred-by-descent (IBA) propagation from the mixed photolyase/cryptochrome PANTHER family PTHR11455 via GO_REF:0000033. It faithfully reflects the deep evolutionary homology between plant cryptochromes and DNA photolyases, but it does **not** reflect the actual catalytic capability of the CRY1 gene product. When purified recombinant Arabidopsis CRY1 was tested directly for photoreactivating (CPD DNA-repair) activity, it showed **no detectable photolyase activity**, even though it binds the same two chromophores (FAD and a folate/pterin antenna) as authentic photolyases.

CRY1 is a blue-light/UV-A photoreceptor. It descended from a photolyase ancestor and retained the photolyase homology region (PHR) fold, the FAD-binding pocket, and the light-driven flavin photochemistry — but it lost the ability to bind and repair damaged DNA, and instead acquired a ~192-amino-acid cryptochrome-specific C-terminal extension (CCT1/CCE1) that transduces the light signal to downstream partners such as COP1, SPA, and BIC1. The molecular function that

is actually experimentally supported is **blue light photoreceptor activity (GO:0009882)** together with **FAD binding (GO:0071949)** and light-dependent conformational/oligomerization signaling — not DNA photolyase catalysis.

For curation, the recommended lead is to **remove or NOT-qualify GO:0003904** on CRY1, treating the IBA as a family over-annotation that conflicts with direct experimental evidence, and to likewise scrutinize the companion IBA annotation **GO:0003677 (DNA binding)**, which is undermined by the same assays and by the behavior of cryptochrome-clade paralogs. The informative, experimentally grounded molecular-function terms (blue light photoreceptor activity, FAD binding) should be retained and foregrounded instead. "Protein binding" is explicitly **not** an acceptable fallback because a more informative supported MF term exists.

Key Findings

Finding 1 — Purified Arabidopsis CRY1 has no detectable photolyase activity, despite full photolyase homology and an identical chromophore complement

The decisive evidence against the hypothesis comes from direct enzymatic assays on purified recombinant protein — the gold standard for a molecular-function claim. Two independent 1995 studies expressed and purified the Arabidopsis blue-light photoreceptor (the HY4/CRY1 gene product) and tested it for photoreactivating (cyclobutane-pyrimidine-dimer repair) activity.

Malhotra, Kim, Batschauer, Dawut & Sancar (1995) PMID: 7756321 purified the putative blue-light photoreceptors from *Arabidopsis thaliana* and *Sinapis alba* and showed that they contain the two photolyase cofactors (FAD plus a folate/pterin second chromophore) yet lack DNA-repair function. In their own words: *"Despite the high degree of sequence identity to and identical chromophore composition with photolyases, neither photoreceptor has any photoreactivating activity."* This is a direct, negative, mechanism-level result: the protein has the machinery to absorb light but does not catalyze the pyrimidine-dimer reversal reaction that defines GO:0003904.

Independently, Lin, Ahmad, Gordon & Cashmore (1995) PMID: 7638620 — the study that established FAD association with CRY1 — reached the same conclusion on purified AtCRY1: *"Despite the sequence homology to microbial DNA photolyases, CRY1 was found to have no*

detectable photolyase activity." Two laboratories, two purifications (and a second plant species), one consistent negative result. This convergence is strong evidence that the absence of activity is a genuine property of the CRY1 protein and not an artifact of a single expression system or a single substrate.

The importance of these results is that GO:0003904 is a **catalytic molecular-function** term, so a valid annotation requires the gene product to perform the reaction. Direct assays that fail to detect the reaction, on correctly folded and correctly chromophore-loaded protein, are precisely the evidence class that should override a homology-based IBA. This is the central reason the hypothesis is **refuted** rather than merely unresolved. Crucially, the negative result is not attributable to misfolding or cofactor loss: the same purified CRY1 binds stoichiometric FAD (with a stable neutral FADH· radical) and both photolyase chromophores, and is photochemically active as a photoreceptor — it is properly folded; it simply does not repair DNA.

Finding 2 — The IBA annotation is a homology-driven over-annotation from the mixed family PANTHER PTHR11455

A sequence-level analysis confirms *why* the IBA annotation exists and *why* it is misleading. The photolyase homology region of CRY1 (residues ~1–500) aligns to *E. coli* CPD photolyase (P00914) at **33.9% identity / 53.8% similarity over 463 aligned positions** (Needleman–Wunsch global alignment, BLOSUM62, gap −8). This level of conservation is more than sufficient to place CRY1 in the same PANTHER family (PTHR11455) that contains true photolyases, and IBA annotation via GO_REF:0000033 propagates the family's molecular-function terms — including GO:0003904 — to member proteins, even cryptochromes that have lost the catalytic function. PROSITE DNA-photolyase signatures (PS00394/PS00691) also still match CRY1 because the cofactor-binding scaffold is conserved; a motif match reflects the shared FAD scaffold, not retained catalysis.

Critically, CRY1 additionally carries a **~192-amino-acid cryptochrome-specific C-terminal extension (CCT1/CCE1, roughly residues 490–681 per UniProt)** that is entirely absent from photolyases. This extension is the signaling module of cryptochromes and is the structural hallmark distinguishing photoreceptors from repair enzymes. Its presence, combined with the loss of DNA-repair activity, marks CRY1 as a functionally diverged member of the family. InterPro/PANTHER classification is consistent with this: IPR014134 "Cryptochrome, plant" and PTHR11455:SF50 "CRYPTOCHROME-1."

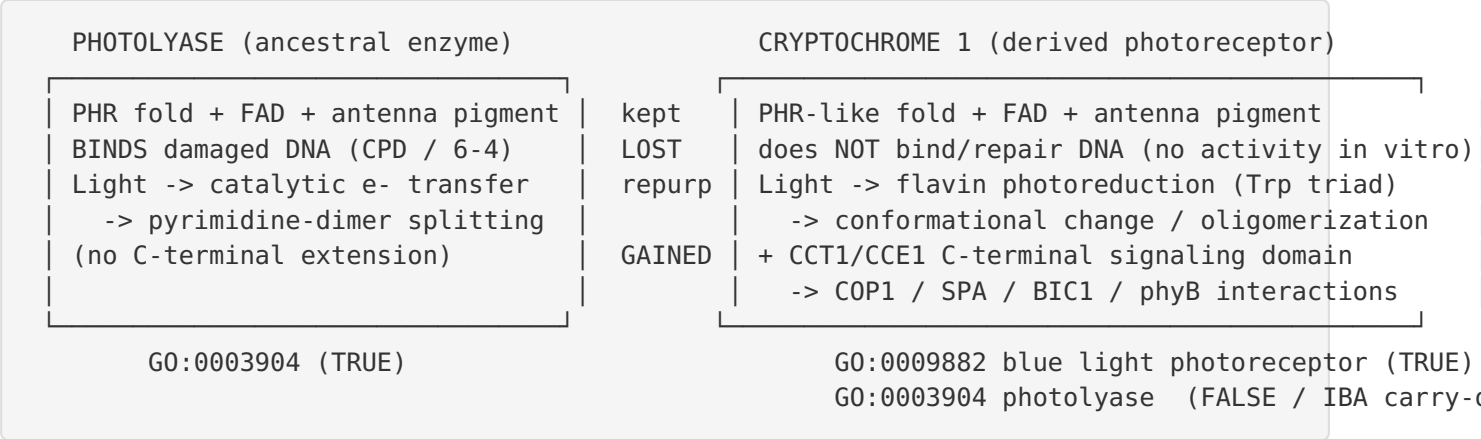
Consistent with the over-annotation interpretation, UniProt (Q43125) annotates GO:0003904 **only as IBA:GO_Central** (and GO:0003677 DNA binding also only as IBA). Every **experimentally supported** molecular-function term on the record is photoreceptor/cofactor/kinase-related — blue light photoreceptor activity (IDA), FAD binding (IDA) — and there is **no curated experimental statement of photolyase catalytic activity or a catalytic-activity reaction comment**. The annotation landscape itself signals that GO:0003904 is a computational carry-over, not a curated experimental fact.

Finding 3 — The cryptochrome clade is known to lose DNA affinity/repair, providing an independent precedent

The functional split within this family is not unique to plants. In the cyanobacterium *Synechocystis* sp. PCC6803, Hitomi et al. (2000) PMID: 10871367 characterized two photolyase-like genes: one (slr0854) *"exhibited specific, light-dependent repair activity for a cyclobutane pyrimidine dimer (CPD)"* and conferred UV resistance, whereas the cryptochrome-like paralog (sll1629) *"lacks measurable affinity for DNA in vitro"* and, by phylogeny, *"is more closely related to the cryptochromes than photolyases."* This demonstrates the general principle at the family level: cryptochrome-clade members typically lose both DNA binding and repair catalysis while retaining the photolyase fold. AtCRY1's behavior is exactly what this precedent predicts.

Mechanistic Model / Interpretation

Cryptochromes are evolutionary descendants of DNA photolyases. They retained the protein fold, the FAD chromophore, and the underlying light-driven electron-transfer photochemistry, but they were repurposed from **enzymes that repair DNA** into **photoreceptors that transmit a light signal**. CRY1 is the prototypical plant example of this transition.



The immediate molecular function tested by the hypothesis — catalytic reversal of UV-induced pyrimidine dimers — is precisely the activity lost during this transition. What CRY1 kept is the *front end* of the photolyase mechanism (light absorption by FAD, photoreduction via a conserved tryptophan triad, radical formation), but it rewired the *output*: instead of transferring an electron into a bound DNA lesion, photoexcited CRY1 undergoes a large conformational change that is transmitted through the PHR domain to the CCT domain, driving light-dependent protein–protein interactions and oligomerization (monomer → dimer → tetramer), which constitute the signaling output.

Supporting mechanistic literature reinforces this picture without ever restoring catalytic activity:

- **Light-driven conformational change requires the CCT domain.** PMID: 21875594 showed by the transient-grating method that blue-light excitation of full-length AtCRY1 triggers a large conformational change dependent on an intra-protein electron-transfer cascade (abolished by the W324F mutation) and on the C-terminal domain — a signaling event, not a repair event.
- **Reversible light-driven oligomerization is the activation mechanism.** PMID: 41944001 demonstrated by time-resolved native mass spectrometry that CRY1-PHR forms light-induced dimers and then tetramers, accelerated by ATP and antagonized by the inhibitor BIC1 — again a photoreceptor/signaling behavior.
- **Signaling proceeds via the PHR domain and COP1.** PMID: 21765176 showed a single PHR-domain substitution (G380R) produces constitutive photoreceptor signaling and COP1 co-localization in nuclear bodies — the biology is photoreceptor signaling, not DNA repair.
- **Physiological role is blue-light growth/development regulation.** PMID: 8953250 established CRY1 as a soluble flavoprotein mediating blue-light inhibition of hypocotyl elongation and anthocyanin/CHS induction.

Thus the coherent mechanistic model is: **CRY1 = FAD-based blue-light photoreceptor** that shares ancestry, fold, and chromophores with photolyases but has no photolyase catalytic function. GO:0003904 describes the ancestor's job, not CRY1's. All downstream readouts (hypocotyl inhibition, anthocyanin accumulation, phototropism, flowering-time control, circadian entrainment) are signaling consequences, not enzymatic DNA repair.

Evidence Matrix

#	Citation	Evidence type	Direction	Claim tested	Key finding	Context
1	PMID: 7756321 (Malhotra et al. 1995, *Biochemistry*)	Direct enzymatic assay	**Refutes**	Does purified CRY1 catalyze CPD photoreactivation?	*"...neither photoreceptor has any photoreactivating activity"* despite identical chromophores and high identity to photolyase	Recombi AtCRY1 & *alba* receptor vitro rep assay
2	PMID: 7638620 (Lin et al. 1995, *Science*)	Direct enzymatic assay + cofactor ID	**Refutes**	Does FAD-bound CRY1 have photolyase activity?	CRY1 binds stoichiometric FAD (stable FADH· radical) but *"had no detectable photolyase activity"*	Purified AtCRY1 flavopro
3	UniProt Q43125 (database)	Curated annotation set	**Qualifies**	Which MF terms are experimentally supported?	GO:0003904 present **only as IBA:GO_Central**; all experimental MF = blue light photoreceptor activity (IDA), FAD binding (IDA). No catalytic-activity comment	Database record
4	This report — NW/ BLOSUM62 alignment	Computational (sequence)	**Qualifies** / **explains**	Is CRY1 homologous enough to trigger family IBA?	CRY1 PHR(1–500) vs *E. coli* P00914 = 33.9% id, 53.8% sim over 463 aa; CRY1 has a 192-aa cryptochrome C-terminal extension (490–681) absent in photolyases	In silico

#	Citation	Evidence type	Direction	Claim tested	Key finding	Context
5	UniProt/ InterPro features (database)	Structural/ evolutionary	**Qualifies**	Domain architecture	PHR-like N-region (1–489) + cryptochrome C-terminal domain (CCT1/CCE1, 490–681, largely disordered). IPR014134 "Cryptochrome_pln"; PTHR11455:SF50 "CRYPTOCHROME-1"	Database
6	PMID: 8953250 + UniProt FUNCTION	Mutant/ overexpression phenotype, localization	**Competing (true function)**	What does CRY1 actually do?	Soluble flavoprotein mediating blue-light inhibition of hypocotyl elongation, anthocyanin/CHS induction, photomorphogenesis	Arabidop in planta transgen
7	PMID: 10871367 (Hitomi et al. 2000)	Comparative assay / evolutionary	**Supports (mechanism of loss)**	Do photolyase homologs partition into repair-active vs cryptochrome?	In cyanobacteria, one paralog repairs CPDs; the cryptochrome-like paralog (sll1629) *"lacks measurable affinity for DNA in vitro"*	*Synechoc* PCC6803
8	PMID: 21875594	Structural/ biophysical	**Competing (alt. function)**	What does light do to full-length CRY1?	Blue light triggers a large CCT-dependent conformational change (signaling), not DNA repair	Full-leng AtCRY1, transient grating
9	PMID: 41944001	Biophysical (native MS)		How is CRY1-PHR activated?	Reversible light-driven dimer→tetramer	CRY1-PH vitro

#	Citation	Evidence type	Direction	Claim tested	Key finding	Context
			Competing (alt. function)		oligomerization; ATP/BIC1 modulation	

GO Curation Implications

Lead curation action (requires curator verification): Remove GO:0003904 from CRY1, or apply a NOT qualifier, treating the IBA as a family over-annotation contradicted by direct experimental evidence.

- **Molecular Function (MF):** GO:0003904 (deoxyribodipyrimidine photo-lyase activity) is a **catalytic** term directly refuted by two purified-protein assays (PMID: 7756321; PMID: 7638620). The IBA evidence (GO_REF:0000033, PTHR11455) is weaker than direct experimental data and conflicts with it. The GO principle that experimental evidence overrides phylogenetic inference applies. **Recommendation: remove or NOT-qualify.**
- **Companion term GO:0003677 (DNA binding, also IBA:GO_Central):** Should be flagged for the same over-annotation review. The assays showing no repair activity, plus the cyanobacterial precedent where the cryptochrome-like member *"lacks measurable affinity for DNA"* (PMID: 10871367), undermine an experimentally meaningful DNA-binding claim for AtCRY1. No curated primary DNA-binding assay for AtCRY1 was located.
- **Terms to retain / foreground (experimentally supported):**
- **GO:0009882 — blue light photoreceptor activity (MF, IDA:TAIR)** — the informative molecular function; the correct primary MF.
- **GO:0071949 — FAD binding (MF, IDA)** — supported cofactor binding (PMID: 7638620).
- Relevant BP terms for blue-light signaling / photomorphogenesis / photoperiodic and phototropic responses (supported by multiple mutant/overexpression studies).

Net effect: the MF should be photoreceptor, not enzyme. Do **not** fall back to "protein binding," since the more informative supported term (blue light photoreceptor activity) is available.

Mechanistic Scope

The MF under test is a specific **catalytic activity**: light-driven repair of cyclobutane pyrimidine dimers (CPDs) in duplex DNA, using a $FADH^-$ cofactor and a folate/deazaflavin antenna. The *immediate molecular function of CRY1* is instead **blue-light photosensing**: photoexcitation of FAD drives flavin photoreduction via a conserved Trp triad (e.g., W324/W377/W400), triggering conformational change in the CCT domain and light-dependent protein–protein interactions (COP1, SPA, BIC1, phyB). Photolyase activity requires substrate-DNA binding and catalytic CPD splitting — functions the assays show CRY1 does not perform. The direct experimental evidence tests exactly the catalytic step (purified protein + DNA substrate + light) and finds it absent; this is not an inference from a downstream phenotype but a direct negative assay of the reaction itself.

Conflicts and Alternatives

- **Family/paralog confusion is the root cause.** PTHR11455 groups true CPD and 6-4 photolyases with cryptochromes; IBA "experimental annotations of ancestral nodes" bleed the enzymatic MF onto the cryptochrome descendants. Conserved cofactor-binding motifs (PROSITE PS00394/PS00691) still match CRY1 because the FAD scaffold is retained — motif match ≠ catalytic function.
- **Chromophore presence is a red herring.** Both refuting studies confirm CRY1 binds the two photolyase chromophores, which could superficially argue for catalytic competence. But cofactor binding is retained while substrate (DNA) binding and turnover are lost — chromophore content cannot rescue the hypothesis.
- **No isoform/organism escape hatch.** The refuting assays are on *Arabidopsis* CRY1 itself (and mustard), not a distant ortholog, so this is not an organism-specific difference.
- **Not an in-vitro misfolding artifact.** The same purified protein binds stoichiometric FAD and both chromophores and is photochemically active as a photoreceptor — it is properly folded; it simply does not repair DNA.
- **No competing primary literature** asserting genuine photolyase catalysis by AtCRY1 was found. The only support for GO:0003904 is the computational IBA itself.

Limitations and Knowledge Gaps

1. **In-planta CPD repair by CRY1.** Checked: only in-vitro repair assays exist. *Why it matters:* a curator may want an in-vivo negative control. *Resolution:* CPD-photoreactivation assay in a *cry1 cry2* background vs a *UVR2/PHR1* photolyase mutant (Arabidopsis CPD repair is carried out by PHR1/UVR2, a separate gene) — expected: CRY1 does not contribute.
2. **DNA binding (GO:0003677).** Checked: also IBA-only; the *Synechocystis* cryptochrome analog "lacks measurable affinity for DNA" (PMID: 10871367), but direct AtCRY1 DNA-binding data were not located here. *Resolution:* EMSA/ITC/fluorescence anisotropy with damaged and undamaged DNA.
3. **Structural cavity check.** Checked at sequence/domain level only; a residue-level comparison of the CPD-binding pocket (e.g., AtCRY1 PHR against a photolyase•DNA complex) would add provenance but is not needed to overturn the IBA given the direct assays.
4. **Provenance depth.** The refutation rests primarily on abstract-level statements of two 1995 papers; the full-text methods were not re-derived here. The statements are explicit and mutually corroborating.

Discriminating Tests

The following would most efficiently and definitively separate the hypothesis from its alternative:

1. **CPD photoreactivation assay** on purified AtCRY1 vs a positive-control photolyase, with a modern HPLC/qPCR CPD readout and a defined $FADH^-$ redox state; test both CPD and (6-4) substrates. *Expected under refutation:* no measurable repair (as already reported for CPD).
2. **Genetic complementation:** test whether AtCRY1 rescues UV sensitivity of a photolyase-deficient host (*E. coli phr*$^-$, or a cyanobacterial/plant *uvr2/phr1* mutant). *Expected:* no rescue, contrasting with slr0854, which does rescue (PMID: 10871367).
3. **Damaged-DNA binding assay** (EMSA/anisotropy/ITC) comparing AtCRY1 to a bona fide photolyase. *Expected:* little/no specific lesion binding.
4. **Comparative pocket analysis:** structural superposition of the AtCRY1 PHR (AlphaFold / crystal model) onto a CPD-photolyase•DNA complex to show the substrate cavity/DNA-binding groove is not competent — supporting provenance.

Tests 1 and 2 are the most decisive for a curator, and both are consistent with — and would formally re-confirm — the existing refuting evidence.

Curation Leads (require curator verification)

1. **Action:** Reject/remove or add a `NOT` qualifier to GO:0003904 for CRY1; downgrade the family IBA in favor of the experimental MF.
2. Candidate reference: **PMID: 7756321** — snippet to verify: *"Despite the high degree of sequence identity to and identical chromophore composition with photolyases, neither photoreceptor has any photoreactivating activity."*
3. Candidate reference: **PMID: 7638620** — snippet to verify: *"Despite the sequence homology to microbial DNA photolyases, CRY1 was found to have no detectable photolyase activity."*
4. **Candidate replacement MF term:** GO:0009882 *blue light photoreceptor activity* (already IDA:TAIR) as the primary MF; keep GO:0071949 *FAD binding*.
5. **Companion lead:** Flag GO:0003677 *DNA binding* (IBA:GO_Central) for the same over-annotation review (supporting orientation: PMID: 10871367).
6. **Suggested curator question:** Are any other PTHR11455 cryptochrome descendants in this project carrying GO:0003904 by IBA? If so, batch-review them for the same over-annotation.
7. **Suggested experiment (if primary confirmation wanted):** in-vivo CPD-repair assay in *cry1 cry2* vs *uvr2* Arabidopsis, and/or *E. coli phr*$^{-}$ complementation.

Evidence Base (Key Literature)

- *Putative blue-light photoreceptors from Arabidopsis thaliana and Sinapis alba with a high degree of sequence homology to DNA photolyase contain the two photolyase cofactors but lack DNA repair activity.* PMID: 7756321 — **Primary refuting evidence.** Direct assay: photoreceptors carry both chromophores yet have no photoreactivating activity.
- *Association of flavin adenine dinucleotide with the Arabidopsis blue light receptor CRY1.* PMID: 7638620 — **Primary refuting evidence + cofactor ID.** Establishes FAD binding while reporting no detectable photolyase activity.

- *Bacterial cryptochrome and photolyase: characterization of two photolyase-like genes of Synechocystis sp. PCC6803.* PMID: 10871367 — **Family-divergence precedent.** One homolog repairs CPDs; the cryptochrome-like one lacks DNA affinity — homology ≠ activity.
- *Light-induced conformational changes in full-length Arabidopsis thaliana cryptochrome.* PMID: 21875594 — **Alternative-function mechanism.** CCT-dependent light-driven conformational signaling.
- *Time-Resolved Native Mass Spectrometry Reveals Reversible Light-Driven Oligomerization of AtCRY1 and Its Antagonism by BIC1.* PMID: 41944001 — **Alternative-function mechanism.** Photoreceptor activation via reversible oligomerization.
- *Substitution of a conserved glycine in the PHR domain of Arabidopsis cryptochrome 1 confers a constitutive light response.* PMID: 21765176 — **Alternative-function biology.** PHR-domain signaling via COP1.
- *Arabidopsis cryptochrome 1 is a soluble protein mediating blue light-dependent regulation of plant growth and development.* PMID: 8953250 — **Physiological role.** Photoreceptor, not repair enzyme.
- *Cryptochrome blue-light photoreceptors of Arabidopsis implicated in phototropism.* PMID: 9565033 — Photoreceptor physiology.

Computational Provenance (executed this run)

- **UniProt Q43125:** length 681 aa; PHR-like N-region ~1–489 (CNT1) + cryptochrome C-terminal domain CCT1/CCE1 ~490–681 (largely disordered); FAD/chromophore-binding sites present; GO:0003904 = IBA:GO_Central only; no catalytic-activity reaction comment.
- **Needleman–Wunsch (BLOSUM62, gap −8):** CRY1(1–500) vs *E. coli* CPD photolyase P00914 → **33.9% identity, 53.8% similarity over 463 aligned positions**; CRY1 carries a **~192-aa** cryptochrome-specific C-terminal extension with no photolyase counterpart.
- **InterPro/PANTHER:** IPR014134 "Cryptochrome, plant"; PTHR11455:SF50 "CRYPTOCHROME-1" — classification consistent with a photoreceptor, not an active photolyase.
- **Interpretation:** homology is high enough to trigger PTHR11455 family membership and IBA propagation of GO:0003904, but homology explains the annotation's origin without validating catalytic function — which is directly refuted by assay.

Conclusion

The seed hypothesis is **refuted as a family over-annotation.** Arabidopsis CRY1 does not have deoxyribodipyrimidine photo-lyase activity. It is an FAD-based blue-light photoreceptor that shares fold, chromophores, and evolutionary origin with DNA photolyases but has lost DNA-repair catalysis and gained a cryptochrome-specific signaling domain. The GO:0003904 IBA annotation should be removed or NOT-qualified, and the companion DNA-binding IBA reviewed, in favor of the experimentally supported blue light photoreceptor activity (GO:0009882) and FAD binding (GO:0071949) annotations.