Cytokinin Biosynthesis in Hexapoda and Insecta:

A Bioinformatic Analysis

Nate Mooi

Scott W. Roy

Edward F. Connor

Department of Biology

San Francisco State University

1600 Holloway Avenue

San Francisco, CA 94132 USA

27 October 2023

Abstract

Cytokinins (CKs) are widespread in a variety of organisms from bacteria to humans, and are particularly abundant in insects. However, how organisms other than bacteria and plants obtain CKs has not been thoroughly studied. We examined the transcriptomes of 670 species of Hexapoda (predominantly Insecta) to determine if transcripts that encode proteins homologous to any of the known enzymes involved in CK biosynthesis and metabolism are widespread in these groups. We found that transcripts encoding proteins homologous to the enzymes tRNA-dimethylallyltransferase (EC: 2.5.1.75) and tRNA-2-methylthio-N6-dimethylallyladenosine synthase (EC: 2.8.4.3) are widespread in the Hexapoda. These enzymes would allow insects and hexapods to synthesize iP-based CKs and methylthiolated iP-based CKs via a tRNA-degradation pathway whereby tRNA is first prenylated and possibly methylthiolated prior to releasing CKs or methylthiolated CKs upon degradation. We also found widespread presence in insects and hexapods of transcripts encoding proteins that are homologous to five enzymes in the adenine salvage pathway: 5'- nucleotidase (EC: 3.1.3.5), adenosine kinase (EC:2.7.1.20), purine-nucleoside phosphorylase (EC: 2.4.2.1), purine nucleosidase (EC: 3.2.2.1), and adenine phosphoribosyltransferase (EC: 2.4.2.7). These enzymes could allow insects and hexapods to convert CK nucleotides to nucleosides and free base CKs. We found few transcripts encoding proteins homologous to enzymes that would convert CKs to storage forms such as their O-glucosides and no transcripts encoding proteins homologous to enzymes that would degrade CKs such as CK oxidases. We suggest that insects and hexapods have the enzymatic pathways necessary to synthesize and metabolize CKs, in contrast to the presumption that CKs are merely obtained via consumption and sequestration from plants or via microbial symbiosis.

Key Words: Cytokinin, biosynthesis, Hexapoda, Insecta, transcriptomes

Introduction

Cytokinins (CK) are a class of N^6-substituted adenine compounds that play a central role in a number of processes including plant growth and development, nutritional signaling, root proliferation, apical dominance, delay of senescence, and shoot meristem function (Mok and Mok 2001, Kamada-Nobusada and Sakakibara 2009). They are considered one of the primary plant hormones. Ironically, the first cytokinin, called kinetin, was discovered and isolated in DNA from the autoclaved sperm of a herring, presumably *Clupea harengus* L. (Miller et al. 1955). For 40 years the presence of this cytokinin, kinetin, in an animal was considered an artifact of extraction and processing, rather than to indicate that forms of the quintessential plant hormone cytokinin also occur endogenously and naturally in animals (Barciszewski et al. 1996).

The narrative describing cytokinin biosynthesis that has developed for plants suggests that the enzyme adenylate dimethylallyltransferase (KEGG - EC: 2.5.1.27 or EC: 2.5.1.112) adds a prenyl group to AMP, ADP, or ATP (adenine mono-, di- or tri-phosphate) to form N^6-isopentenyladenosine-5'- mono-, di-, or tri-phosphate. It is thought that this process is responsible for the bulk of cytokinin biosynthesis in plants (Klámbt 1992, Sakakibara 2006, Spíchal 2012, Kieber and Schaller 2014). In plant plastids, the methylerythritol pathway (MEP), and in the cytosol, the mevalonate pathway (MVA), are responsible for the synthesis of the prenyl group donors dimethylallyl pyrophosphate (DMAPP) and 4-hydroxy-3-methyl-2-but-3-enyl-pyrophosphate (HMBDP) (Tarkowská and Strnad 2018).

A second cytokinin biosynthesis pathway has been documented in plants and involves the prenylation of a tRNA bound adenine molecule at position A 37 of the anti-codon loop (tRNA-dimethylallyltransferase; EC: 2.5.1.75) which after hydroxylation of the prenyl group and upon

degradation of the tRNA results in the release of *cis*-zeatin (*c*Z) (Konevega et al. 2006, Chimnaronk et al. 2009, Dabravolski 2020). However, the tRNA based pathway has been considered unproductive, at least in part because tRNA turnover has been considered to be insufficient to support the bulk of CK biosynthesis in most plant species (Klámbt 1992, Sakakibara 2006, Spíchal 2012, Kieber and Schaller 2014). Furthermore, no enzyme for the conversion of *c*Z to *trans*-zeatin (*t*Z) has been isolated, and in most plant species forms of *t*Z and isopentenyladenine (iP) are the most abundant forms of CK (Gajdošová et al. 2011). The prenylation of tRNA at position A 37 is thought to stabilize codon-anticodon interactions to ensure fidelity of translation (Ubonovičius et al. 2001, Miyawaki *et al*. 2006, Agris et al. 2007,). In plants, further modification of the prenylated tRNA by methylthiolation leads to the production of the methylthiolated CK, 2-methylthio-N6-(*cis*-hydroxyisopentenyl) adenosine (ms^2i^6A) upon tRNA degradation (Dabravolski 2020, Gibb et al. 2020) (Figure 1).

A number of enzymes in the adenine salvage pathway would be capable of removing phosphate groups and ribose from nucleotide and nucleoside forms of CK hence generating nucleoside and freebase forms of these CKs (Galuszka et al. 2008, Frebort et al. 2011, Ashihara et al. 2018, Figure 2). However, in plants the LOG gene (EC: 3.2.2.-, cytokinin riboside 5'-monophosphate phosphoribohydrolase) can efficiently convert nucleotide forms of iP, *t*Z, and *c*Z to freebase forms which are considered the active forms of CKs (Kurakawa et al. 2007, Kuroha et al. 2009).

Soon after the discovery of CKs in plants, CKs or compounds with CK-like activity were reported in gall-inducing and leaf-mining insects (McCalla et al. 1962, Engelbrecht 1968, Engelbrecht et al. 1969, Ohkawa 1974, van Staden 1975). However, only very recently have CKs been reported in insects other than gall-inducing and leaf-mining species (Brütting et al. 2018,

Andreas et a*l*. 2020, Tokuda et al. 2022). CKs have now been detected in animal groups besides insects, having been reported in dogs, humans, nematodes, and broadly in the Arthropoda (De Lillo and Montfreda 2004, Siddique et al. 2015, Seegobin et al. 2018, Aoki et al. 2019, Tokuda et al. 2022). CKs are also found in bacteria (Armstrong et al. 1969, Akiyoshi et al. 1987), fungi (Hinsch et al. 2015, Morrison et al. 2015a,b, Chanclud et al. 2016), algae (Lu and Xu 2015, Žižková et al. 2017), and protists (Ludwig-Müller et al. 2009, Malinowski et al. 2016).

Morrison et al. (2015a) suggest that, in fungi, CK biosynthesis must involve the tRNA-degradation pathway since the adenylate dimethylallyltransferase pathway is absent. By similar reasoning it would appear that in other groups, except a limited number of bacterial species, CK biosynthesis would also occur via the tRNA-degradation pathway, since the adenylate dimethylallyltransferase gene (EC: 2.5.1.27, EC: 2.5.1.112) appears to be restricted to plants, several genera of bacteria (e.g., *Agrobacterium*, *Pantoe*a, *Erwinia*, *Streptomyces*, *Xanthomonas*, *Ralstonia*, *Pseudomonas*, etc.), and the slime mold *Dictyostelium* (Linder et al. 2014, Nishii et al. 2018, Wang et al. 2020, Andreas et al. 2020). Morrison et al. (2015a) also suggest that adenine salvage enzymes would be capable of inter-converting nucleotide and nucleoside forms of CKs to active freebase forms in fungi.

Andreas et al. (2020) performed a comprehensive analysis of CK profiles for seventeen insect species that were either gall-inducing or their non-gall-inducing close relatives. They report considerable variation among species and among life-history stages within species in their CK profiles, but that nucleotide, nucleoside, and freebase forms of *t*Z, iP and *c*Z were most common. In some species methylthiolated CKs and CK O-glucosides were also present in lower concentrations.

How then do insects obtain CKs? Three potential routes by which insects might acquire CKs include: 1) consumption and sequestration from their diet, 2) via a bacterial or fungal symbiont that synthesizes CKs, and 3) biosynthesis by the insects themselves. The whole body and gland-specific concentrations of CKs in insects often far exceeds that for CKs reported in plant tissues (Mapes and Davies 2001, Yamaguchi et al. 2012, Andreas et al. 2020, Jia et al. 2020,Tokuda et al. 2022). Based on these concentration differences and the small quantity of plant tissue consumed by early stage juvenile insects, Straka et al. (2010), Bartlett and Connor (2013), and Andreas et al. (2020) argue that the concentrations of CKs observed in insects cannot be the result of consumption and sequestration. The possibility of microbial symbiosis as a source of CKs remains a viable option in specific instances, however, several lines of evidence argue against symbiotic acquisition of CKs as a widespread mechanism in insects. Hammer et al. (2021) examined several gall-inducing and non-gall-inducing, but closely related insects using bacterial 16S ribosomal RNA gene amplicon analysis and found no evidence for a shared symbiont, among the gall-inducing species, or for a community of symbionts that characterized the gall-inducing guild. While Andreas et al. (2020) reported that CKs are also found in non-gall inducing insects, there was no suggestion in the data that species with an obligate intra-cellular symbiont, or species with an extensive microbiome had higher CK concentrations than those without symbionts. Ponce et al. (2021) used antibody staining to localize CKs to the salivary glands in the gall-inducing fly, *Eurosta solidaginis* (Diptera: Tephritidae). However, a DNA specific stain did not reveal any bacterial chromosomes in the salivary glands, suggesting that the *Wolbachia* endosymbiont of *E. solidaginis w*Esol is not associated with CK production. Examination of the genome of *w*Esol indicates that it is compromised in the synthesis of DMAPP and S-adenosyl L-methionine (SAM, involved in methylthiolation of CKs), as are the

majority of strains of *Wolbachia* (Fiutek et al. 2023). Hence, the currently available evidence argues against microbial symbiosis and therefore it is most likely that insects acquire CKs via biosynthesis (Andreas et al. 2020).

To determine if the enzymatic machinery necessary for CK biosynthesis and inter-conversion to active freebase forms is widespread in insects, we used a bioinformatic approach to examine the transcriptomes of 670 insect and hexapod species. We sought evidence for enzymes to synthesize CKs and methylthiolated CKs, enzymes in the adenine salvage pathway to inter-convert CKs to active freebase forms, enzymes to form CK O-glucosides, and enzymes to catabolize CKs.

Methods

Our goal was to search for transcripts that encode homologs of enzymes in any of the currently described pathways for the biosynthesis and metabolism of CKs in a large sample of insect and hexapod species. We were particularly interested in determining if homologs to enzymes involved in CK biosynthesis and metabolism are present in a wide array of insects and hexapods. A tRNA-degradation pathway has been proposed for the synthesis of CKs in fungi and insects (Morrison et al. 2015a, Andreas et al. 2020, Fig. 1). The adenine salvage pathway has been proposed for inter-conversion of nucleotides, nucleosides, and freebase forms of CKs (Sakakibara 2006, Galuszka et al. 2008, Frébort et al. 2011, Morrison et al. 2015a, Figure 2).

Transcriptomes and Enzymes

We used all insect and hexapod transcriptomes generated as part of the 1Kite Project, an ongoing effort to amass the transcriptomes of 1000 arthropod species. As of August 30, 2021, the

database contained 670 unique transcriptomes of insect and hexapod species distributed across many insect and hexapod orders. We classified the habitat and trophic mode for juvenile and adult individuals of each hexapod species. We also recorded the life stage responsible for the RNA sample used to construct each transcriptome (Supplemental Table 1).

We used the pathway map for zeatin biosynthesis (00908) from the Kyoto Encyclopedia of Genes and Genomes (KEGG) to visualize and outline the potential pathways for CK biosynthesis, as well as to gather exemplars of the enzymes involved in these pathways (Kanehisa and Goto 2000, Kanehisa 2019, Kanehisa et al. 2021). We also used enzymes from the part of the purine metabolism pathway map (00230) that are associated with the inter-conversion of adenine and AMP. We included animal homologs of the *miaB* gene (CDK5RAP1, EC: 2.8.4.3), and exemplars for the LOG gene (EC: 3.2.2.-). The KEGG EC and KO numbers for each exemplar enzyme are given in Table 1. The amino acid sequences of the complete exemplar proteins were compiled into a Diamond BLAST searchable database.

Our list of target enzymes contained 17 enzymes from cytokinin biosynthesis and metabolism and adenine salvage enzymes (Table 1). Three additional enzymes are hypothesized, but no exemplars were available, so we could not search for them. These three enzymes include cytokinin *trans*-hydroxylase to convert iP to *t*Z (no KEGG EC number, but KO K10717), zeatin reductase to convert *t*Z to dihydro-zeatin (EC:1.3.1.69), and *cis*-hydrolase to convert iP or *t*Z to *c*Z (hypothesized only). Cytokinin *trans*-hydrolase has been identified in plants and is a P450 monooxygenase. P450 monooxygenases are a large family of enzymes that is well represented in insects (Gould 1984, Feyereisen 1999, Scott and Wen 2001). Rather than risk identifying many non-homologous and non-orthologous transcripts, we chose not to use the available plant exemplars and did not search for this enzyme.

183 To search for transcripts encoding proteins homologous to enzymes in the CK and
184 adenine salvage pathways, we established a two step process based on BLAST alignments, first
185 to a set of exemplars, and second, for those with alignments, to a set of insect genomes. We
186 aligned each insect transcriptome against our database of KEGG exemplars using Diamond
187 BLASTx (Buchfink et al. 2015). Our exemplar database consisted of over 500 proteins listed in
188 KEGG as homologs of our target enzymes (Supplemental Table 2). To identify transcripts
189 encoding proteins homologous to these exemplars in our transcriptomes, we retained transcripts
190 with BLAST alignments with e-values below 1 x $10e^{-10}$. To ensure that we were not simply
191 identifying paralogs, we next used annotated insect genomes to exclude transcripts that had a
192 best BLAST alignment to a gene annotated as an enzyme other than the CK-related or adenine
193 salvage enzymes of interest. To accomplish this, we established a BLAST searchable database
194 consisting of the proteins from a set of 181 genomes annotated using the NCBI eukaryotic
195 genome annotation pipeline (Supplemental Table 3). Our database included all available NCBI
196 annotated insect genomes (167), though we supplemented the database with several vertebrate
197 and plant genomes as well. We used Diamond BLASTx to align the transcripts identified
198 through the KEGG-exemplar search described above against this set of annotated genomes and
199 retained transcripts that encoded proteins as potential homologs if their best BLAST alignment
200 (e-values below 1 x $10e^{-10}$) was to a protein annotated as one of our candidate proteins. We used
201 annotations included in the KEGG description of each enzyme as well as the refseq or GenBank
202 annotations in the headers of the protein files for each exemplar as synonyms. Some enzymes
203 had several monikers, all of which we took to be synonyms. This second step involved both an
204 assessment of sequence similarity and functional annotation as a basis of retaining transcripts
205 that encode proteins as potential homologs for enzymes in CK pathways. Annotations interpreted

as indications that a specific enzyme had been detected are provided in Supplemental Table 4. Some annotations lead us to reject specific alignments as paralogs (Supplemental Table 5). This two-step pipeline first narrowed the set of transcripts to be examined further via BLASTx against our KEGG exemplar enzyme database, and then compared this reduced set of transcripts against a set of well-annotated genomes for both sequence similarity and identity in functional annotation to enzymes in CK biosynthesis and adenine salvage pathways. A diagram outlining this two-step pipeline to identify transcripts encoding proteins in insect transcriptomes homologous to enzymes in CK biosynthesis and adenine salvage pathways is depicted in Figure 3.

Examination of protein alignments annotated as "uncharacterized protein"

The results of the BLAST alignments of the KEGG-identified transcripts that encoded proteins with homology to CK and adenine salvage pathway enzymes against the database of annotated genomes produced best alignments for many transcripts to "uncharacterized protein." This is not surprising, as "uncharacterized protein" annotations are common even among well-studied genes in poorly studied genomes. Even some of the enzymes we used from KEGG that were presented as homologs of enzymes for CK biosynthesis were annotated in other databases such as refseq or Genbank as "uncharacterized protein." Because the updating of annotations in NCBI temporally lags the knowledge of protein function, for all transcripts that aligned to an "uncharacterized protein," we collected these proteins and used BLASTp with default settings to align them to the NCBI non-redundant protein data base (NR) and examined the top 25 alignments for each, excluding self-alignments. For those transcripts encoding proteins that had a clear functional annotation to a CK or adenine salvage enzyme as the second best alignment,

we report these transcripts as encoding protein homologs, but list these alignments as alignments of secondary quality. We clearly indicate in which analyses we combine these secondary alignments with transcripts identified initially to encode proteins homologous to enzymes in CK biosynthesis and adenine salvage pathways.

Estimating false negative and false positive protein detection rates

To determine if using the proteins encoded by the transcripts we detected with the two-step pipeline outlined above overlooked some protein homologs or tended to detect some non-homologous proteins, we examined a random sample of 60 transcriptomes to estimate our error rates. We aligned the entire transcriptome of each species in our sample set using BLASTx against our database of 181 well-annotated genomes collected from NCBI. We assigned this set of sample proteins encoded by transcripts detected to be the "truth" and compared the results obtained from our pipeline that first used alignments against the set of KEGG exemplar enzymes. Proteins encoded by transcripts detected in the whole transcriptome BLASTs that were not detected in the KEGG exemplar BLASTs were considered false negatives, and proteins encoded by transcripts detected in the KEGG exemplar BLASTs that were not detected in the whole transcriptome BLASTs were considered false positives. We then expressed our false positive rate as the number of proteins encoded by transcripts only detected in our two-step pipeline divided by the number of proteins encoded by transcripts detected in the whole transcriptome alignments and our false negative rate as the number of proteins encoded by transcripts not detected in our two-step pipeline divided by the number of proteins encoded by transcripts detected in the whole transcriptome alignments. Our estimated error rates also ignored any potential homolog whose top BLAST alignment was annotated as "uncharacterized protein",

252 which was discussed in the previous section. Finally, we tallied the number of transcriptomes in
253 which we detected transcripts encoding protein homologs for each enzyme.
254
255 **Results**
256 From the 1Kite project we used the transcriptomes of 670 species from Insecta and
257 Hexapoda (Supplemental Table 1). Ten species had two transcriptomes each of similar quality in
258 the 1Kite dataset, but we randomly selected a single transcriptome to represent each species.
259 Our set of transcriptomes spanned 31 orders and 374 families of Hexapoda and Insecta and was
260 much more diverse phylogenetically compared to the sequenced genomes available for insects
261 (e.g, our reference set of NCBI annotated insect genomes;167 species, and the full set of
262 annotated insect genomes; 493 species). Note that among all insect genomes sequenced, Diptera
263 and Hymenoptera combined comprise 59.2%, among NCBI annotated genomes Diptera and
264 Hymenoptera combined comprise 71.4%, but among the 1Kite transcriptomes Diptera and
265 Hymenoptera combined comprise only 20.3%. The biases in the phylogenetic distribution of the
266 dataset from the 1Kite project and the genome datasets are clearly apparent (Figure 4). The
267 majority of species in our sample were terrestrial in habitat (Supplemental Figure 1), and our
268 sample included ten trophic modes including herbivorous, mycophagous, predaceous, parasitic,
269 omnivorous, scavenger, and detritivorous species (Supplemental Figure 2). For the
270 overwhelming majority of species, RNA was obtained from the adult stage (Supplemental Figure
271 3).
272 Based on our two-step pipeline built to identify homologs of CK-related and adenine
273 salvage enzymes, taking care to exclude potential paralogs, examination of a random sample of
274 60 insect transcriptomes indicated that our two-step pipeline yielded very low rates of false

positive and false negative alignments. Our overall false positive rate was of 1.0 % and false negative rate was 0.67%. Our highest per enzyme error rate was a false positive rate of 3.3% for adenosine kinase (EC:2.7.1.20, our enzyme e9). Only four enzymes had non-zero error rates.

Our transcriptome-wide search detected homologs of enzymes for CK biosynthesis and adenine salvage as being widespread in insects and hexapods (Figure 5 and Supplemental Table 1). Approximately 83% of transcriptomes yielded evidence that insects and hexapods possess homologs of the enzymes necessary to synthesize iP-based CKs and to interconvert AMP to adenosine, and adenine. Among enzymes that lead to the synthesis of CKs either directly or via the tRNA-degradation pathway, we detected two enzymes as being widespread among insects and hexapods. We found transcripts that encoded homologs of the tRNA-dimethylallyltransferase enzyme (EC 2.5.1.75, our enzyme e2) in 82.9% of the insect and hexapod transcriptomes. This enzyme modifies tRNA and upon degradation of the tRNA leads to the release of CKs. We also identified transcripts that encoded homologs of the CDK5RAP1 enzyme (EC: 2.8.4.3, our enzyme e3) in 83.2% of the transcriptomes. CDK5RAP1 modifies prenylated tRNAs and ultimately leads to the release of methylthiolated CKs upon tRNA degradation. For transcripts that encoded proteins homologous to tRNA-dimethylallyltransferase or CDK5RAP1, 75% or more species had best alignments with e-values $< 10^{-30}$ (see Supplemental Table 6). However, in agreement with Andreas et al. (2020) we found almost no evidence of an adenylate dimethylallyltransferase enzyme (EC: 2.5.1.27, EC 2.5.1.112, our enzyme e1), which leads directly to the synthesis of CKs from adenine, in any insect or hexapod. In only a single species, *Chrysura cuprea* (Hymenoptera: Chrysididae), did we find a transcript that encoded a protein with a best alignment to an adenylate dimethylallyltransferase enzyme. However, when aligned to the entire NCBI non-redundant protein database (NR) this transcript-

encoded protein best alignment was to a tRNA-dimethylallyltransferase enzyme (EC 2.5.1.75, our enzyme e2), not to an adenylate dimethylallyltransferase enzyme. We also did not detect homologs of the MiaE enzyme (EC: 1.14.99.69, our enzyme e4), which leads to the release of methylthiolated *c*Z, in any of the insect or hexapod transcriptomes.

In the adenine salvage pathway, we found transcripts that encoded proteins with significant alignments to four enzymes in almost all insect and hexapod transcriptomes (Figure 5 and Supplemental Table 1). Transcripts encoding proteins homologous to two of these enzymes, 5'-nucleotidase (EC: 3.1.3.5, our enzyme e5) and adenosine kinase (EC: 2.7.1.20, our enzyme e9), both involved in converting adenine nucleotides to nucleosides (e.g., removing phosphate from AMP to produce adenosine), were detected in 99.1% and 92.8% of the transcriptomes, respectively. We also found significant alignments of transcripts encoding proteins to the enzyme purine-nucleoside phosphorylase (EC: 2.4.2.1, our enzyme e8), which is involved in removing the ribose from adenosine to generate the freebase iP, in 92.1% of the insects and hexapod transcriptomes. Together these enzymes are capable of converting AMP to adenine. For all three of these enzymes, more than 75% of species had transcripts encoding proteins with best BLAST alignments with e-values $<10^{-100}$ (Supplemental Table 6). Transcripts encoding proteins homologous to the enzyme adenine phosphoribosyltransferase (EC: 2.4.2.7, our enzyme e10), which directly converts AMP to adenine, were detected in 87.6% of the transcriptomes with over 80% of species having best BLAST alignments with e-values $<10^{-30}$ (Supplemental Table 3). Finally, transcripts that encoded proteins with significant alignments to another enzyme capable of performing the conversion of adenosine to adenine, purine nucleosidase (EC: 3.2.2.1, our enzyme e6), were also detected in almost half of the transcriptomes (44.9%). However, for two other enzymes capable of directly converting AMP to adenine, we found few or no species with

transcripts encoding proteins with significant alignments. For the LOG gene, cytokinin riboside 5'-monophosphate phosphoribohydrolase (EC: 3.2.2.-, our enzyme e7), we found a few transcripts encoding proteins with significant alignments scattered throughout the Insecta. For the enzyme AMP nucleosidase (EC: 3.2.2.4, our enzyme e11), we found no transcripts encoding proteins with significant alignments in any insect or hexapod transcriptome.

Of the enzymes involved in CK inactivation whether through the formation of O-glucosides or the complete degradation of CKs, we found few or no transcripts encoding proteins with significant alignments (Figure 5 and Supplemental Table 1). Of the glycosyltransferases, only zeatin O-glucosyltransferase (EC: 2.4.1.203, our enzyme e12) had any transcripts that encoded proteins with best BLAST alignments (Figure 5). We found no transcripts encoding proteins with best BLAST alignments to the CK degrading cytokinin oxidase (EC: 1.5.99.12, our enzyme e17).

Examination of the transcripts that in the second step of our two-step pipeline annotated as "uncharacterized protein" revealed that most of these transcripts did not encode any of our target enzymes. However, we did find a large number of transcripts that encoded proteins that were annotated as "uncharacterized protein" that had second best alignments to the enzyme purine nucleosidase (EC: 3.2.2.1, our enzyme e6). When we combined the primary alignment and the secondary alignments for this enzyme, the percentage of species with at least one transcript encoding a protein that align to purine nucleosidase increased from 44.9% to 83.9% (Supplemental Table 7) and >75% of species had best BLAST alignments with e-values< 10^{-30} (Supplemental Table 4).

Discussion

Given the importance of cytokinins (CK) in plant biology and in the corresponding use of CKs by plant-associated insects to manipulate plants, the source of CKs used by insects is of great interest. In this work, we searched 670 insect and other hexapod transcriptomes for CK-related enzymes. Contrary to widespread assumptions, we find that CK-related enzymes are ubiquitous in insects and hexapods, including lineages without known close associations with plants. This pattern suggests that CK production and metabolism is common in and ancestral to insects, and thus that insects manipulate plants by co-opting pre-existing innate CK-producing pathways in the insects themselves, rather than by consumption and sequestration of plant-derived CKs or stimulation of CK production by plants, as has often been assumed.

Limitations of Bioinformatic Approach

A bioinformatic approach such as ours is limited as a basis for confirming the enzymes involved in CK biosynthesis and metabolism in insects in a number of ways: 1) we are restricted to examining only enzymes for which exemplars are available, 2) we must assume that the functions of the exemplar enzymes are as annotated, and 3) we must treat absence of a transcript that encodes a homolog of an enzyme from a transcriptome as evidence merely of non-detection, rather than absence of that enzyme from the genome. If insects have evolved a non-canonical means to synthesize or metabolize CKs then our approach will be unable to detect it. If enzymes have functions other than those annotated, or are capable of using substrates other than those expected, then we are likely to miss enzymes critical to the biosynthesis and metabolism of CKs. The inefficiencies of RNA extraction, reverse complementation, library preparation, sequencing, assembly, and stage and tissue specific expression of genes will almost guarantee that a

transcriptome will not represent the entire genome. Ten species among our sample of 670 had two transcriptomes each in the 1KITE dataset. Among the seventeen enzymes we examined, a few were present in only a single transcriptome within a pair for one or more of the reasons mentioned above. Therefore, we will always underestimate the breadth of distribution of each enzyme. Given these caveats, enzymes for which we detected transcripts that encoded proteins homologous to the exemplar enzymes in virtually no species are likely not present in insects and hexapods. Enzymes for which we detected transcripts that encoded proteins that were homologous to the exemplar enzymes in the vast majority of species are probably more widespread than our estimates indicate. Bias in the set of KEGG exemplars and in our comparison database of NCBI annotated genomes toward the Holometabola could also make it more difficult to detect transcripts encoding proteins homologous to enzymes for CK biosynthesis and metabolism in hexapod groups outside the Holometabola since they are under-represented in the exemplars and genome database.

The low error rates we estimated in the comparison of our two-step analysis pipeline to whole transcriptome alignments suggest that our approach is sound and computationally efficient when facing limited computing resources. Furthermore, we only detected a few false positive and negative errors, which trivially add or subtract from our estimation of the breadth of distribution of CK biosynthesis and metabolism enzymes.

Cytokinin Biosynthesis in Hexapoda

Our examination of the transcriptomes of 670 species of Hexapoda suggest that the ability to synthesize CKs and methylthiolated CKs via the tRNA degradation pathway is widespread in these groups. Furthermore, we also detected transcripts encoding proteins

390 homologous to enzymes capable of the inter-conversion of AMP to adenine, which could convert
391 CK-nucleotides to CK-nucleosides to freebase forms of CK. Finally, we found no evidence of
392 widespread enzymes in hexapods for the conversion of CKs into storage forms such as O-
393 glucosides, and no evidence for a homolog of a CK oxidase enzyme capable of permanently
394 degrading CKs. Furthermore, based on a graphical examination of the frequencies of detecting
395 each enzyme (analysis not shown), we found no tendency for the detection of any of these
396 enzymes to depend on the habitat in which the species is found, or on the trophic mode of
397 juveniles or adults.

398 Earlier ideas that CKs in insects were derived from consumption and sequestration or
399 microbial symbiosis were in part motivated by the narrative surrounding CK biosynthesis that
400 had developed for plants. CKs were and still are considered phytohormones, so their occurrence
401 in insects and other animals was thought to be from an exogenous source (Giron et al. 2013,
402 Bartlett and Connor 2014). How could animals possibly be synthesizing an important plant
403 hormone? However, the widespread distribution of the tRNA-degradation pathway for CK
404 biosynthesis we report here, coupled with the widespread distribution and abundance of CKs in
405 insects argues for CK biosynthesis by insects (Andreas et al. 2020, Tokuda et al. 2022).
406 Furthermore, Chen et al. (2019) examined the BmIPT1 gene from *Bombyx mori* (Lepidoptera:
407 Bombycidae) and showed that it is a functional tRNA-dimethylallyltransferase gene since it
408 could complement a mutant yeast strain lacking tRNA-dimethylallytransferase. Using HPLC,
409 Chen et al. (2019) detected isopentenyladenosine in the recombinant yeast strain complemented
410 by BmIPT1, but isopentenyladenosine was absent from the mutant yeast. Finally, the MVA
411 pathway for the biosynthesis of the prenyl group donor DMAPP and the enzymes for
412 biosynthesis of the methyl group donor S-adenosyl L-methionine (SAM) from dietary

413 methionine (EC: 2.5.1.6) are also widespread in insects (Bellés et al. 2004, Hayashi et al. 2022).
414 Hence, insects can produce the substrates necessary to synthesize CKs and methylthiolated CKs.
415 Contrary to the narrative characterizing cytokinin biosynthesis in plants, CK biosynthesis
416 in animals appears to exclusively employ the tRNA-degradation pathway as evidenced by the
417 lack of any transcripts encoding proteins homologs of adenylate dimethylallyltransferase
418 enzymes having been detected in our examination of the transcriptomes of 670 species in the
419 Hexapoda or reported in any other animal group (Lindner et al. 2014, Nishii et al. 2018, Wang et
420 al. 2020, Andreas et al. 2020). Furthermore, in contrast to plants where the tRNA-degradation
421 pathway leads predominantly to the production of *c*Z and only low levels of iP (Sakakibara 2006,
422 Gibb et al. 2020), in animals it leads exclusively to the production of iP (Persson et al. 1994).
423 Forms of iP and *t*Z have been found to be the most abundant CKs in insects (Andreas et al. 2020,
424 Tokuda et al. 2022). Finally, the productivity of the tRNA-degradation pathway which has been
425 questioned for plants (Klámbt 1992, Spíchal 2012, Kieber and Schaller 2014), appears to be
426 capable of generating much higher concentrations of CKs in animals than in plants (Andreas et
427 al. 2020). Perhaps the higher concentration of CKs in animals arises because of higher turnover
428 rates of tRNA or because the tRNA-dimethylallyltransferase enzyme is more promiscuous in
429 animals. In a careful study partitioning the total pool of RNA in HeLa cells into tRNA, poly-A-
430 RNA, miRNA, and rRNA pools, Reiter et al. (2012) found abundant prenylated tRNA, miRNA,
431 and rRNA, similar abundances of methylthiolated forms in these pools, and a small amount of
432 prenylated poly-A-RNA. They also found abundant methylthiolated poly-A-RNA, which implies
433 that most of the prenylated poly-A-RNA is rapidly methylthiolated, since RNAs must first be
434 prenylated before they can be methylthiolated. Reiter et al.'s (2012) experiments suggest that
435 tRNA-dimethylallyltransferase and tRNA-2-methylthio-N6-dimethylallyladenosine synthase

enzymes (CDK5RAP1) are capable of modifying a wide range of RNA species which upon degradation release CKs, and this may also be true in plants as well. Hence, it would appear that the tRNA-degradation pathway, which should potentially be renamed the RNA-degradation pathway, is more than capable of generating high concentrations of CKs contrary to earlier views stemming from work exclusively in plants.

In retrospect, the presence of the tRNA-degradation pathway for CK biosynthesis in animals should be no surprise. tRNA modifications are widespread and the specific tRNA modifications that lead to CK biosynthesis and methylthiolated CK biosynthesis are ancestral, occurring in all domains of life (Anton et al. 2008, Čavužić and Liu 2017, Wang et al. 2020, Dabravolski 2020), in spite of earlier claims that the tRNA-dimethylallyltransferase enzyme was absent from the Archaea (Frebort et al. 2011).

Adenine Salvage Pathway in Hexapoda

Enzymes for the metabolism of adenine are present in all domains of life so the fact that we found transcripts encoding proteins homologous to five enzymes involved in inter-conversion of AMP, adenosine, and adenine to be widespread in Hexapoda should also be no surprise (Armenta-Medina et al. 2014). Presumably these enzymes have at least some affinity for adenine molecules that have been prenylated or methylthiolated as has been reported for these enzymes from plants (Sakakibara 2006, Galuszka et al. 2008, Frébort et al. 2011). Therefore they would be capable of removing phosphate and riboside groups from isopentenyladenosine phosphate to yield active, freebase iP. However, recent work suggests that ribosides of CKs also have physiological effects in plants and are not simply inactive precursors of the freebase forms (Nguyen et al. 2021).

Other Cytokinin Metabolism Enzymes

We had anticipated that few transcripts homologous to glycosyltransferases or CK oxidases would be detected in our analysis of hexapod transcriptomes. Given the primary role of the tRNA-dimethylallytransferase in ensuring accurate protein translation, we conjectured that hexapods would have no enzymes to regulate the pool of CKs or methylthiolated CKs via conversion to temporary storage forms such as O-glucosides or complete degradation of CKs by a CK oxidase as do plants. While the expression of the tRNA-dependent enzymes may be amplified to generate higher concentrations of CKs to be secreted for manipulation of host plants (Yamaguchi et al. 2014, Brütting et al. 2018, Ponce et al. 2022), the CK pool would not be finely regulated as in plants, but rather via secretion during feeding or oviposition. In hexapods, CKs do not appear to serve as signaling molecules as they do in plants.

Coda

The bioinformatic analysis we performed supports the idea that hexapods synthesize CKs and methylthiolated CKs using an RNA-dependent degradation pathway. This pathway results initially in the synthesis of iP nucleotides, as well as methylthiolated forms. This is consistent with the available evidence that shows that forms of iP are among the most abundant forms of CKs in insects and hexapods (Andreas et al. 2020 and Tokuda et al. 2022), and that methylthiolated CKs are also present in insects (Andreas et al. 2020). Enzymes in the adenine salvage pathway are also present and likely involved in the inter-conversion of forms of iP nucleotides into nucleosides and to freebase isopentenyladenine and 2-methylthio-

isopentyladenine. These same enzymes are also likely involved in the inter-conversion of other forms of CK nucleosides into freebase forms.

Although we did not attempt to parse the large number of P450 monooxygenase enzymes found in insects to generate candidate enzymes capable of inter-converting iP and *t*Z, the high concentrations of *t*Z observed in insects suggest that such an enzyme is present and critical to the high concentrations of CKs observed in insects (Andreas et al. 2020, Tokuda et al. 2022). Andreas et al. (2020) detected CK-glucosides in a subset of the species they examined, but almost always in low concentrations which is consistent with the general absence of enzymes for the synthesis of CK-glucosides that we report here. Enzymes for the synthesis of CK-glucosides could be present in only a subset of species, or potentially CK-glucosides could be formed non-enzymatically. Finally, we found no evidence for a CK-oxidase in hexapods, which suggests that CK pools may be regulated in insects and hexapods solely by a balance of biosynthesis and secretion.

In hexapods, endogenous cytokinins appear to function to improve fidelity of protein translation and in the manipulation of food plants. Determining if CKs have other functions in hexapods will require more observations and careful experiments.

Acknowledgements

We would like to thank Dr. José R. de la Torre for his guidance and computational assistance throughout this research. We also thank Tomáš Furmánek and Daniel Johnson from the SFSU Department of Academic Technology for their timely configuration and maintenance of our virtual server to support our computation. We also thank Natalie Fiutek, Stephannie Seng,

and Joshua Natahusada for their support and encouragement during this period of pandemic impacted research.

Statements and Declarations

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Nate Mooi and Edward F. Connor. The first draft of the manuscript was written by Nate Mooi and Edward F. Connor and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

The authors declare no conflicts of interest.

Literature Cited

Agris PF, Vendeix FAP, Graham WD (2007) tRNA's wobble decoding of the genome:40 years of modification. J Mol Biol 366: 1-13. https://doi.org/10.1016/j.jmb.2006.11.046

Akiyoshi DE, Regier DA, Gordon MP (1987) Cytokinin production by *Agrobacterium* and *Pseudomonas* spp. J Bacteriol 169: 4242-4248. https://doi.org/10.1128/jb.169.9.4242-4248

Andreas P, Kisiala A, Emery RJN, de Clerck-Floate RM,Tooker JF, Price PW, Miller DG III, Chen MS, Connor EF(2020) Cytokinins are abundant and widespread among insect species. Plants 9(2): 208. https:/doi.org/10.3390/plants9020208

Anton BP, Saleh L, Benner JS, Raleigh EA, Kasif S, Roberts RJ (2008) RimO, a MiaB-like enzyme, methylthiolates the universally conserved Asp88 residue of ribosomal protein S12 in *Escherichia coli.* Proc Natl Acad Sci USA. 105: 1826-1831. https://doi.org/10.1073pnas.0708608105

526 Aoki M, Seegobin M, Kisiala A, Noble A, Brunetti C, Emery RJN (2019) Phytohormone
527 metabolism in human cells: Cytokinins are taken up and interconverted in HeLa cell
528 culture. FASEB BioAdv 1: 320-331. https://doi.org/10.1096/fba.2018-00032
529 Armenta-Medina D, Segovia L, Perez-Rueda1 E (2014) Comparative genomics of nucleotide
530 metabolism: a tour to the past of the three cellular domains of life. BMC Genomics 15(1):
531 800. https://doi.org/10.1186/1471-2164-15-800
532 Armstrong D J, Burrows W J, Skoog E, Roy K L, Soli D (1969) Cytokinins: Distribution in
533 transfer RNA species of *Escherichia coli*. Proc Natl Acad Sci USA 63:834-41.
534 https://doi.org/10.1073/pnas.63.3.834
535 Ashihara H, Stasolla C, Fujimura T, Crozier A (2018) Purine salvage in plants. Phytochem 147:
536 89e124. https://doi.org/10.1016/j.phytochem.2017.12.008
537 Barciszewski J, Rattan SIS, Siboska G, Clark BFC (1999) Kinetin - 45 years on. Plant Sci 148:
538 37-45. https://doi.org/10.1016/S0168-9452(99)00116-8
539 Bartlett L, Connor EF (2014) Exogenous phytohormones and the induction of plant galls by
540 insects. Arthropod-Plant Interact 8: 339-348. https://doi.org/10.1007/s11829-014-9309-0
541 Bellés X, Martín D, Piulachs M-D (2005) The mevalonate pathway and the synthesis of juvenile
542 hormone in insects. Ann Rev Entomol 50: 181-199. https://doi.org/
543 10.1146/annurev.ento.50.071803.130356
544 Brütting C, Crava CM, Schäfer M, Schuman MC, Meldau S, Adam N, Baldwin IT (2018)
545 Cytokinin transfer by a free-living mirid to *Nicotiana attenuata* recapitulates a strategy of
546 endophytic insects. eLife 7:e36268. https://doi.org/10.7554/eLife.36268
547 Buchfink B, Xie C, Huson D H (2015) Fast and sensitive protein alignment using Diamond. Nat
548 Methods 12: 59-60. https://doi.org/10.1038/nmeth.3176

549 Čavužić M, Liu Y (2017) Biosynthesis of sulfur-containing tRNA modifications: a comparison
550 of bacterial, archaeal, and eukaryotic pathways. Biomolecules 7: 27.
551 https://doi/org/10.3390/biom7010027
552 Chanclud E, Kisiala A, Emery RJN, Chalvon V, Ducasse A, Romiti-Michel C, Gravot A, Kroj T,
553 Morel J (2016) Cytokinin production by the rice blast fungus is a pivotal requirement for
554 full virulence. PloS Pathog 12(2): e1005457.
555 https://doi.org/10.1371/journal.ppat.1005457
556 Chen Y, Bai B, Yan H, Wen F, Qin D, Jander G, Xia Q, Wang G (2019) Systemic disruption of
557 the homeostasis of transfer RNA isopentenyltransferase causes growth and development
558 abnormalities in *Bombyx mori*. Insect Mol Biol 28: 380–391.
559 https://doi.org/10.1111/imb.12561
560 Chimnaronk S, Forouhar F, Sakai J, Yao M, Tron C.M, Atta M, Fontecave M, Hunt JF, Tanaka I
561 (2009) Snapshots of dynamics in synthesizing N6- isopentenyladenosine at the tRNA
562 anticodon. Biochem 48: 5057-5065. https://doi.org/10.1021/bi900337d
563 De Lillo E, Monfreda R (2004) 'Salivary secretions' of eriophyoids (Acari: Eriophyoidea): first
564 results of an experimental model. Exp Appl Acarol 34: 291–306.
565 https://doi.org/10.1007/s10493-004-0267-6
566 Dabravolski S (2020) Multi-faceted nature of the tRNA isopentenyltransferase. Funct Plant Biol
567 47: 475–485. https://doi.org/10.1071/FP19255
568 Engelbrecht L (1968) Cytokinin in den "grünen inseln" des herbstlaubes. Flora 159: 208-214.
569 https://doi.org/10.1016/S0367-1836(17)30103-9
570 Engelbrecht L, Orban U, Hesse W (1969) Leaf-miner caterpillars and cytokinins in the "green
571 islands" of autumn leaves. Nature 223: 319-321. https://doi.org/10.1038/223319a0

572 Feyereisen R (1999) Insect P450 Enzymes. Ann Rev Entomol 44: 507-533.
573 https://doi.org/10.1146/annurev.ento.44.1.507
574 Fiutek N, Couger MB, Roy SW, de la Torre JR, Connor EF (2023) Genomic assessment of the
575 contribution of the *Wolbachia* endosymbiont of *Eurosta solidaginis* to gall-induction . Int
576 J Mol Sci 24: 9613. https://doi.org/10.3390/ijms24119613
577 Frébort I, Kowalska M, Hluska T, Frébortová J, Galuszka T (2011) Evolution of cytokinin
578 biosynthesis and degradation. J Exp Bot 62: 2431-2452.
579 https://doi.org/10.1093/jxb/err004
580 Gajdošová S, Spíchal L, Kamínek M, Hoyerová K, Novák O, Dobrev PI, Galuszka P, Klíma P,
581 Gaudinová A, Žižková E, Hanuš J, Dančák M, Trávniček B, Pešek B, Krupička M,
582 Vaňková R, Strnad M, Motyka V (2011) Distribution, biological activities, metabolism,
583 and conceivable function of *cis*-zeatin-type cytokinins in plants. J Exp Bot 62: 2287-
584 2840. https://doi.org/10.1093/jxb/erq457
585 Galuszka P, Spíchal L, Kopečný D, Tarkowski P, Frébortová J, Šebela M, Frébort I (2008)
586 Metabolism of plant hormones cytokinins and their functions signaling, cell
587 differentiation and plant development. Stud Nat Prod Chem 34: 203-263.
588 https://doi.org/10.1016/S1572-5995(08)80028-2
589 Gibb M, Kisiala AB, Morrison EN, Emery RJN (2020) The origins and roles of methylthiolated
590 cytokinins: Evidence from among life kingdoms. Front Cell Dev Biol 8: 605672.
591 https://doi.org/10.3389/fcell.2020.605672
592 Giron D, Frago E, Glevarec G, Pieterse CMJ, Dicke M (2013) Cytokinins as key regulators in
593 plant–microbe–insect interactions: connecting plant growth and defence. Funct Ecol 27:
594 599–609. https://doi.org/10.1111/1365-2435.12042

595 Gould F (1984) Mixed function oxidases and herbivore polyphagy: the devil's advocate position.
596 Ecol Entomol 9:29–34. https://doi.org/10.1111/j.1365-2311.1984.tb00695.x
597 Hammer TJ, De Clerck-Floate R, Tooker JF, Price PW, Miller DG III, Connor, EF (2021) Are
598 bacterial symbionts associated with gall induction in insects? Arthropod-Plant Interact
599 15: 1-12. https://doi.org/10.1007/s11829-020-09800-6
600 Hayashi Y, Kashio S, Murotomi K, Hino S, Kang W, Miyado K, Nakao M, Miura M, Kobayashi
601 S, Namihira M (2022) Biosynthesis of S-adenosyl-methionine enhances aging-related
602 defects in *Drosophila* oogenesis. Sci Rep 12: 5593. https://doi.org/10.1038/s41598-022-
603 09424-1
604 Hinsch J, Vrabka J, Oeser B, Novák O, Galuszka P, Tudzynsji P (2015) De novo biosynthesis of
605 cytokinins in the biotrophic fungus *Claviceps purpurea*. Environ Microbiol 17: 2935-
606 2951. https://doi.org/10.1111/1462-2920.12838
607 Jia M, Li Q, Hua J, Liu J, Zhou W, Qu B and Luo S (2020) Phytohormones regulate both "fish
608 scale" galls and cones on *Picea koraiensis*. Front Plant Sci 11: 580155. https://doi.org/
609 10.3389/fpls.2020.580155
610 Kamada-Nobusada T, Sakakibara H (2009) The molecular basis of cytokinin biosynthesis.
611 Phytochem 70: 444-449. https:/doi.org/10.1016/j.phytochem.2009.02.007
612 Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids
613 Res 28: 27-30. https://doi.org/10.1093/nar/28.1.27
614 Kanehisa M (2019) Toward understanding the origin and evolution of cellular organisms. Protein
615 Sci 28: 1947-1951. https://doi.org/10.1002/pro.3715

616 Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M (2021) KEGG: integrating
617 viruses and cellular organisms. Nucleic Acids Res 49: D545-D551.
618 https://doi.org/10.1093/nar/gkaa970
619 Kieber, J.J. and G.E. Schaller (2014) Cytokinins. The Arabidopsis Book 2014: e0168. https://
620 doi.org/10.1199/tab.0168
621 Klámbt, D. (1992) The biogenesis of cytokinins in higher plants: Our present knowledge. In
622 Physiology and Biochemistry of Cytokinins in Plants; Kamínek M, Mok DWS,
623 Zažímalová E, Eds.; SPB Academic Publishing: The Hague, The Netherlands; pp. 25–27.
624 ISBN 978-9051030662.
625 Konevega AL, Soboleva NG, Makhno VI, Peshekhonov AV, Katunin VI (2006) The effect of
626 modification of tRNA nucleotide-37 on the tRNA interaction with the P- and the A-site
627 of the 70s ribosome *Escherichia coli*. Mol Biol 40: 669-683.
628 https://doi.org/10.1134/S0026893306040121
629 Kurakawa T, Ueda N, Maekawa M, Kobayashi K, Kojima M, Nagato Y, Sakakibara H, Kyozuka
630 J (2007) Direct control of shoot meristem activity by a cytokinin-activating enzyme.
631 Nature 445: 652-655. https://doi.org/10.1038/nature05504
632 Kuroha T, Tokunaga H, Kojima M, Ueda N, Ishida T, Nagawa S, Fukuda H, Sugimoto K,
633 Sakakibara H (2009) Functional analyses of LONELY GUY cytokinin-activating
634 enzymes reveal the importance of the direct activation pathway in Arabidopsis. Plant Cell
635 21: 3152–3169. https:// doi. org/ 10. 1105/ tpc. 109. 068676
636 Lindner AC, Lang D, Seifert M, Podlešáková K, Novák O, Strnad M, Reski R, von
637 Schwartzenberg K (2014) Isopentenyltransferase-1 (IPT1) knockout in *Physcomitrella*

638 together with phylogenetic analyses of IPTs provide insights into evolution of plant
639 cytokinin biosynthesis. J Exp Bot 65: 2533–2543. https://doi.org/10.1093/jxb/eru142
640 Ludwig-Müller J, Prinsen E, Rolfe SA, Scholes JD (2009) Metabolism and plant hormone action
641 during clubroot disease. J Plant Growth Regul 28: 229–244.
642 https://doi.org/10.1007/s00344-009-9089-4
643 Lu Y, Xu J (2015) Phytohormones in microalgae: a new opportunity for microalgal
644 biotechnology? Trends Plant Sci 20: 273-282.
645 http://dx.doi.org/10.1016/j.tplants.2015.01.006
646 Malinowski R, Novák O, Borhan MH, Spíchal L, Strnad M, Rolfe SA (2016) The role of
647 cytokinins in clubroot disease. Eur J Plant Pathol 145: 543–557.
648 https://doi.org/10.1007/s10658-015-0845-y
649 Mapes CC, Davies PJ (2001) Cytokinins in the ball gall of *Solidago altissima* and the gall
650 forming larvae of *Eurosta solidaginis.* New Phytol 151: 203-212. https://doi.org/
651 10.1046/j.1469-8137.2001.00158.x
652 McCalla DR, Genthe MK, Hovanitz W (1962) The chemical nature of an insect gall-growth
653 factor. Plant Physiol 37: 98-103. https://doi.org/10.1104/pp.37.1.98
654 Miller CO, Skoog F, Von Saltza MH, Strong FM (1955) Kinetin, a cell division factor from
655 deoxyribonucleic acid. Science 77: 1392. https://doi.org/10.1021/ja01610a105
656 Miyawaki K, Tarkowski P, Matsumoto-Kitano M, Kato T, Sato S, Tarkowski D, Tabala S,
657 Sanberg G & Kakimoto T (2006) Roles of *Arabidopsis* ATP/ADP isopentenyltransferases
658 and tRNA isopentenyltransferases in cytokinin biosynthesis. Proc Natl Acad Sci USA
659 103: 16598-16603. https://doi.org/10.1073/pnas.0603522103

660 Mok DWS and Mok MC (2001) Cytokinin metabolism and action. Annu Rev Plant Physiol Plant
661 Mol Biol 52: 89-118. https://doi.org/10.1146/annurev.arplant.52.1.89
662 Morrison EN, Knowles S, Thorn RG, Saville BJ, Emery RJN (2015) Detection of
663 phytohormones in temperate forest fungi predicts consistent abscisic acid production and
664 a common pathway for cytokinin biosynthesis. Mycologia 107: 245-257.
665 https:/doi.org/10.3852/14-157
666 Morrison EN, Emery RJN, Saville BJ (2015) Phytohormone involvement in the *Ustilago maydis*-
667 *Zea mays* pathosystem: relationships between abscisic acid and cytokinin levels and
668 strain virulence in infected cob tissue. PLoS ONE 10(6): e0130945.
669 https://doi.org/10.1371/journal.pone.0130945
670 Nguyen HN, Nguyen TQ, Kisiala AB, Emery RJN (2021) Beyond transport: cytokinin ribosides
671 are translocated and active in regulating the development and environmental responses of
672 plants. Planta 254: 45. https://doi.org/10.1007/s00425-021-03693-2
673 Nishii K, Wright F, Chen YY, Möller M (2018) Tangled history of a multigene family: The
674 evolution of *ISOPENTENYLTRANSFERASE* genes. PLoS ONE 13(8): e0201198.
675 https:/doi.org/10.1371/journal.pone.0201198
676 Ohkawa M (1974) Isolation of zeatin from larvae of *Drycocosmus kuriplilus* Yasumatsu.
677 Hortscience 9: 458-459.
678 Persson BC, Esberg B, Ólafsson Ó, Björk GR (1994) Synthesis and function of
679 isopentenyladenosine derivatives in tRNA. Biochimie 76: 1152-1160.
680 https://doi.org/10.1016/0300-9084(94)90044-2

Ponce GE, Fuse M, Chan A, Connor EF (2021) The localization of phytohormones within the gall-inducing insect *Eurosta solidaginis* (Diptera: Tephritidae). Arthropod-Plant Interact 15: 375-385. https://doi.org/10.1007/s11829-021-09817-5

Reiter V, Matschal DMS, Wagner M, Globisch D, Kneuttinger AC, Muller M, Carell T (2012) The CDK5 repressor CDK5RAP1 is a methylthiotransferase acting on nuclear and mitochondrial RNA Nucleic Acids Res 40: 6235-6240. https://doi.org/1093/nar/gks240

Sakakibara H (2006) Cytokinins: activity, biosynthesis, and translocation. Annu Rev Plant Biol 57: 431–449. https://doi.org/10.1146/annurev.arplant.57.032905.105231

Scott JG, Wen Z (2001) Cytochromes P450 of insects: the tip of the iceberg. Pest Manag Sci 57: 958-967. https://doi.org/10.1002/ps.354

Spíchal, L. (2012) Cytokinins – recent news and views of evolutionally old molecules. Funct Plant Biol 39L 267-284. http://dx.doi.org/10.1071/FP11276

Seegobin M, Kisiala A, Noble A, Kaplan D, Brunetti C, Emery RJN (2018) *Canis familiaris* tissues are characterized by different profiles of cytokinins typical of the tRNA degradation pathway. FASEB J 32: 6575-6581. https://doi.org/10.1096/fj.201800347

Siddique S, Radakovic ZS, De La Torre CM, Chronis D, Novák O, Ramireddy E, Holbein J, Matera C, Hütten M, Gutbrod P, Anjam MS, Rozanska E, Habash S, Elashry A, Sobczak M, Kakimoto T, Strnad M, Schmülling T, Mitchum MG, Grundler FMW (2015) A parasitic nematode releases cytokinin that controls cell division and orchestrates feeding site formation in host plants. Proc Natl Acad Sci USA 112: 12669-12674. https://doi.org/10.1073/pnas.1503657112

Straka J, Hayward A, Emery R (2010) Gall-inducing *Pachypsylla celtidis* (Psyllidae) infiltrate hackberry trees with high concentrations of phytohormones. J Plant Interact 5: 197-203. https://doi.org/10.1080/17429145.2010.484552

Tarkowská D, Strnad M (2018) Isoprenoid-derived plant signaling molecules: biosynthesis and biological importance. Planta 247: 1051-106. https://doi.org/10.1007/s00425-018-2878-x

Tokuda M, Suzuki Y, Fujita S, Matsuda H, Adachi-Fukunaga S, Elsayed AK (2022) Terrestrial arthropods broadly possess endogenous phytohormones auxin and cytokinins. Sci Rep 12: 4750. https://doi.org/10.1038/s41598-022-08558-6

Ubonovičius J, Qian Q, Durand, JMB, Hagervall TG, Björk GR (2001) Improvement of reading frame maintenance is a common function for several tRNA modifications. EMBO J 20: 4863-4873. https://doi.org/10.1093/emboj/20.17.4863

van Staden J (1975) Cytokinins from larvae in *Erythrina latissima* galls. Plant Sci Lett 5: 227-230. https://doi.org/10.1016/0304-4211(75)90016-4

Wang X, Lin S, Liu D, Gan L, McAvoy R, Ding J, Li Y (2020) Evolution and roles of cytokinin genes in angiosperms 1: Do ancient IPTs play housekeeping while non-ancient IPTs play regulatory roles? Horticulture Res 7: 28. https://doi.org/10.1038/s41438-019-0211-x

Yamaguchi H, Tanaka H, Hasegawa M, Tokuda M, Asami T, Suzuki Y (2012) Phytohormones and willow gall induction by a gall-inducing sawfly. New Phytol 196: 586-595. https://doi.org/ 10.1111/j.1469-8137.2012.04264.x

Žižková E, Kubeš M, Dobrev PI, Přibyl P, Šimura J, Zahajská L, Drábková LZ, Novák O, Motyka V. 2017. Control of cytokinin and auxin homeostasis in cyanobacteria and algae. Ann Bot 119: 151-166. https://doi.org/10.1093/aob/mcw194

Table 1. Potential cytokinin biosynthesis and adenine salvage enzymes

Enzyme Function	KEGG EC Number	KEGG KO Number	Enzyme	Abbreviation	Our number
cytokinin biosynthesis	EC: 2.5.1.27, EC:2.5.1.112	K10760	adenylate dimethylallyltransferase	IPT	e1
tRNA modifying	EC: 2.5.1.75	K00791	tRNA dimethylallyltransferase	TRIT1, miaA	e2
	EC: 2.8.4.3	K06168	tRNA-2-methylthio-N6-dimethylallyladenosine synthase	CDK5RAP1, miaB	e3
	EC:1.14.99.69	K06169	*t*RNA 2-(methylsulfanyl)-N6-isopentenyladenosine37 hydroxylase	miaE	e4
Adenine Salvage	EC: 3.1.3.5	K01081, K19970, K03787, K11751, K08693	5'- nucleotidase		e5
	EC: 3.2.2.1	K01239	purine nucleosidase		e6
	EC: 3.2.2.-	K22522	cytokinin riboside 5'-monophosphate phosphoribohydrolase	LOG	e7
	EC:2.4.2.1	K03783, K03784, K09913	purine-nucleoside phosphorylase	punA	e8
	EC:2.7.1.20	K00856	adenosine kinase	ADK	e9
	EC:2.4.2.7	K00759	adenine phosphoribosyltransferase	APRT	e10
	EC:3.2.2.4	K01241	AMP nucleosidase	amn	e11
formation of O- and N-glucosides of cytokinins	EC:2.4.1.203	K13492	zeatin O-glucosyltransferase	ZOG1	e12
	EC:2.4.1.215	K13495	*cis*-zeatin O-glucosyltransferase	CISZOG	e13
	EC:2.4.1.-	K13496, K23452	UDP-glucosyltransferase 73C & 85A	UGT73C UGT85A	e14
	EC:2.4.2.40	K13494	zeatin O-xylosyltransferase	ZOX1	e15
	EC:2.4.1.-	K13493	cytokinin-N-glucosyltransferase	UGT76C1_2	e16
inactivate CKs	EC:1.5.99.12	K00279-7	cytokinin dehydrogenase/oxidase	CKX	e17

Figure Captions

Figure 1. CK and methylthiolated CK biosynthesis pathway (Dabravolski 2020). Beginning with a tRNA that reads codons starting with U and have adenine at position 36, 37, and 38, the enzyme EC: 2.5.1.75 (tRNA-dimethylallyltransferase) transfers a prenyl group from DMAPP (dimethylallyl pyrophosphate) to form i^6A37 (isopentenyl adenosine embedded in the tRNA). After the tRNA degrades, the nucleotide, isopentenyl adenosine monophosphate is released (iPNT). Prior to tRNA degradation, a second enzymatic transformation in which the i^6A37 tRNA is methylthiolated involves the enzyme EC: 2.8.4.3 (tRNA-2-methylthio-N6-dimethylallyladenosine synthase), S-Adenosyl L-methionine, and a thiol group which forms ms^2i^6A37 embedded in the tRNA. Upon tRNA degradation, the nucleotide, 2MeSiPNT (2-methylthio-isopentenyladenosine monophosphate) is released.

Figure 2. Enzymes in the adenine salvage pathway, along with the LOG gene (EC:3.2.2.-) projected to be involved in conversion between adenine nucleotides (AMP), adenine nucleosides (adenosine), and adenine (see Table 1). Bidirectional enzymatic reactions are shown in dark blue, and unidirectional enzymatic reactions in light blue. Enzymes determined to be widespread in insects and hexapods and depicted using solid lines and those not detected are depicted as dashed lines. For purine nucleosidase only (e6, EC:3.2.2.1), the determination that the enzyme is widespread involved combining the transcripts initially identified to encode proteins homologous to purine nucleosidase in

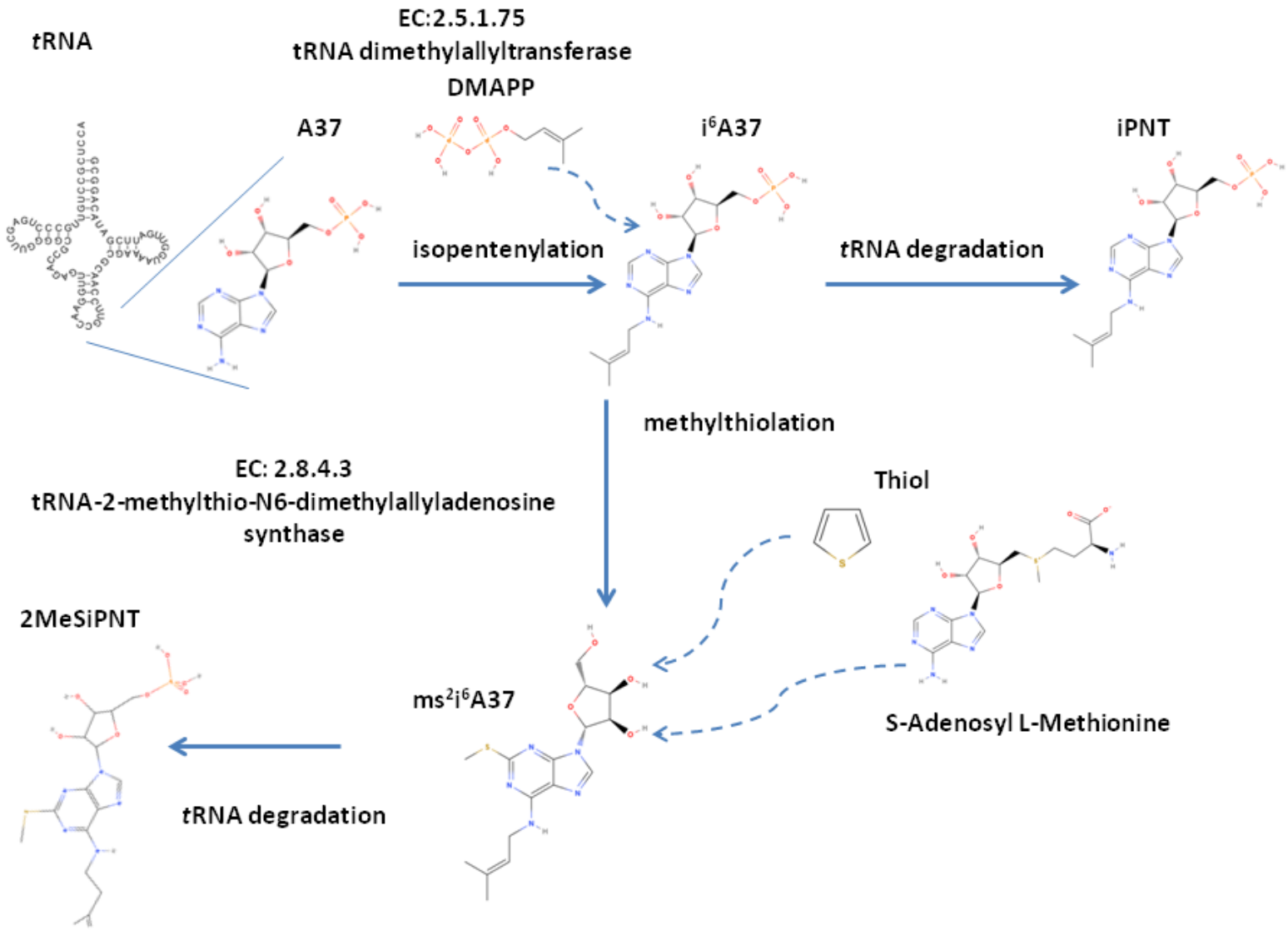
combination with transcripts initially identified as "uncharcaterized protein" that when aligned to NR using BLASTp had second best alignments to purine nucleosidase.

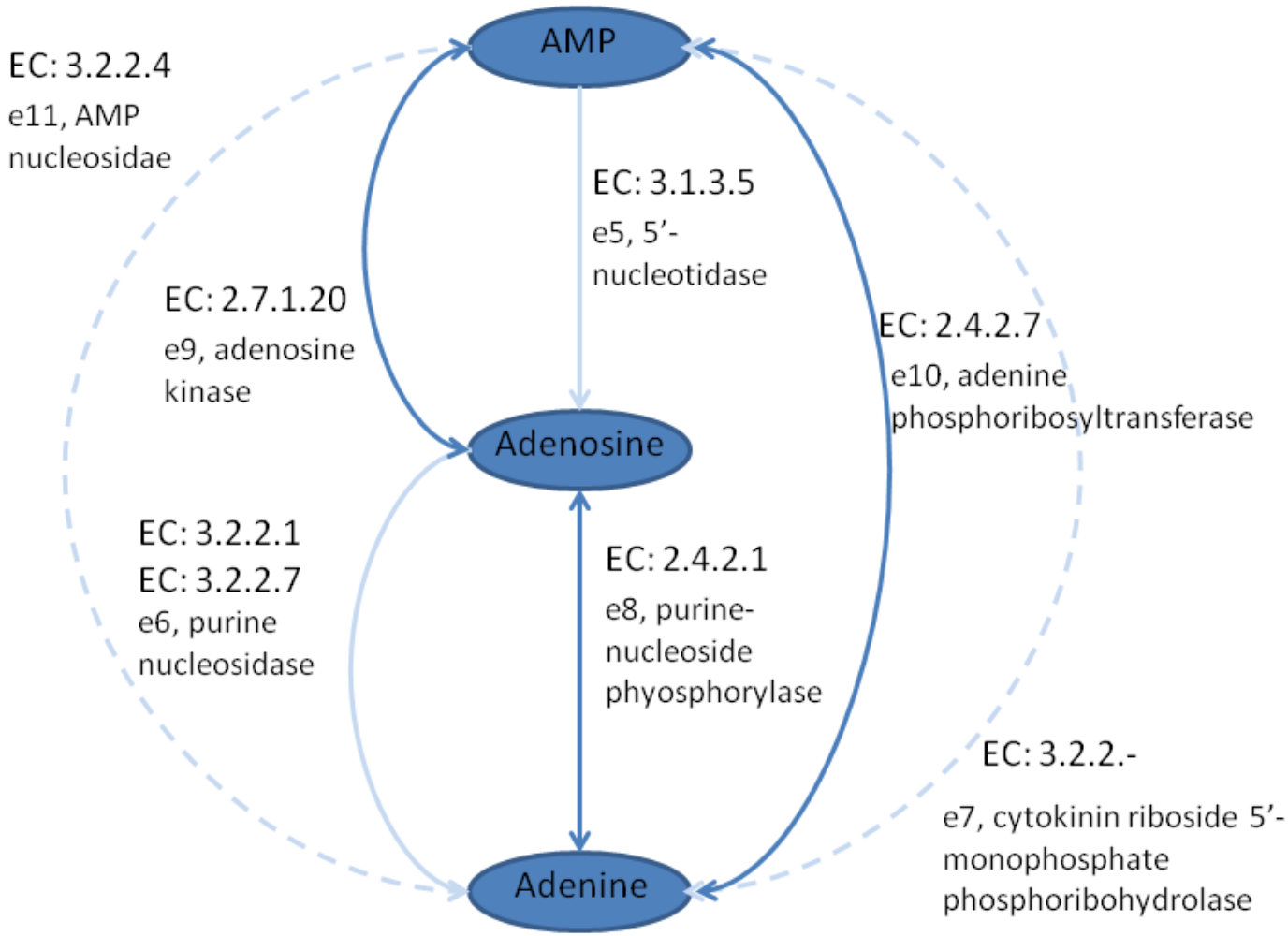
Figure 3. Two-step bioinformatic pipeline to identify transcripts that are homologous to enzymes in CK and adenine salvage pathways. Blue vertical line separates Step 1 from Step 2.

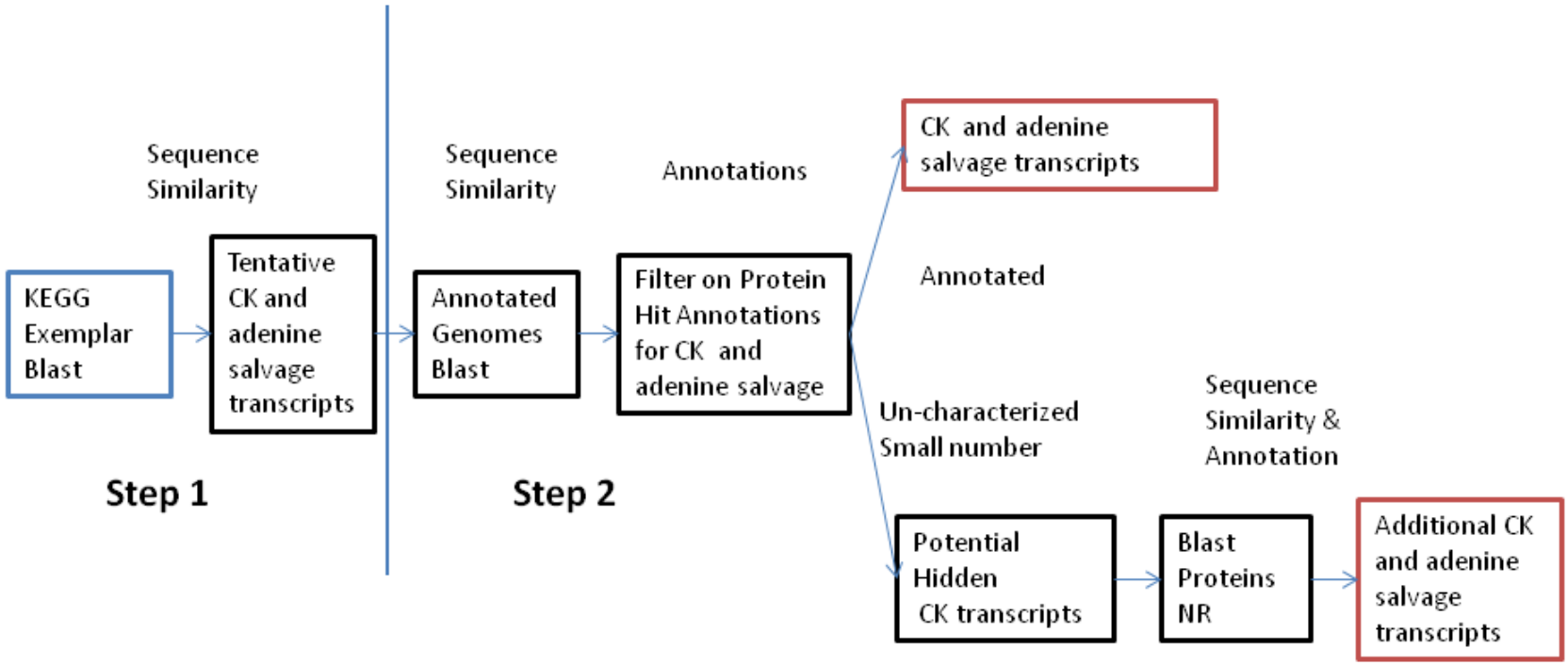
Figure 4. Representation of insect species among orders in our sample of 1Kite transcriptomes (670 species) in comparison to the representation among all insect genomes (493 species) and among the insect genomes annotated by the NCBI annotation pipeline (167 species).

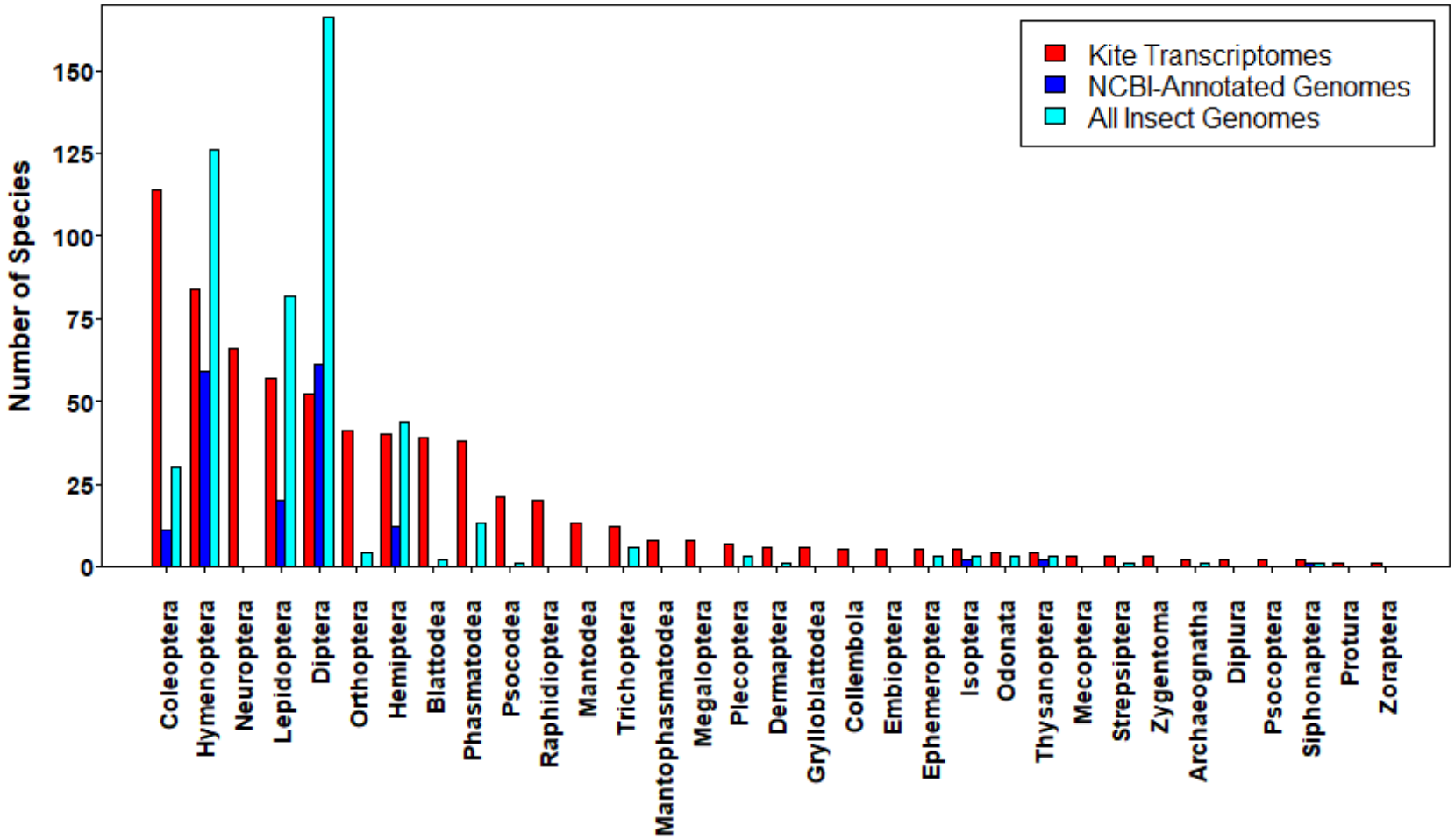
Figure 5. Heat map showing pattern of detection of homologs of enzymes involved in known CK biosynthesis and adaneine salvage pathways in transcriptomes of 670 species of insect or hexapod. The heat map depicts the proportion of species in each of 374 families of hexapods and insects in which at least one homolog of each enzyme was detected. Each section of the multicolored bar on the right side of the heat map is proportional to the number of families in each order that were sampled. Darker green colors indicate higher proportions of species possess homologs. Enzymes are divided into four functional groups by the blue vertical bars: 1) CK Synthesis - includes both de novo synthesis and tRNA modification enzymes, 2) Adenine salvage enzymes, 3) Glucosyltransferases involved in synthesis of O- and N-glucosides of CKs, and conversion of *tZ* to dihydrozeatin, and 4) Cytokinin Oxidase which inactivates CKs. Note that enzymes for tRNA modification and adenine salvage are widespread in insects and

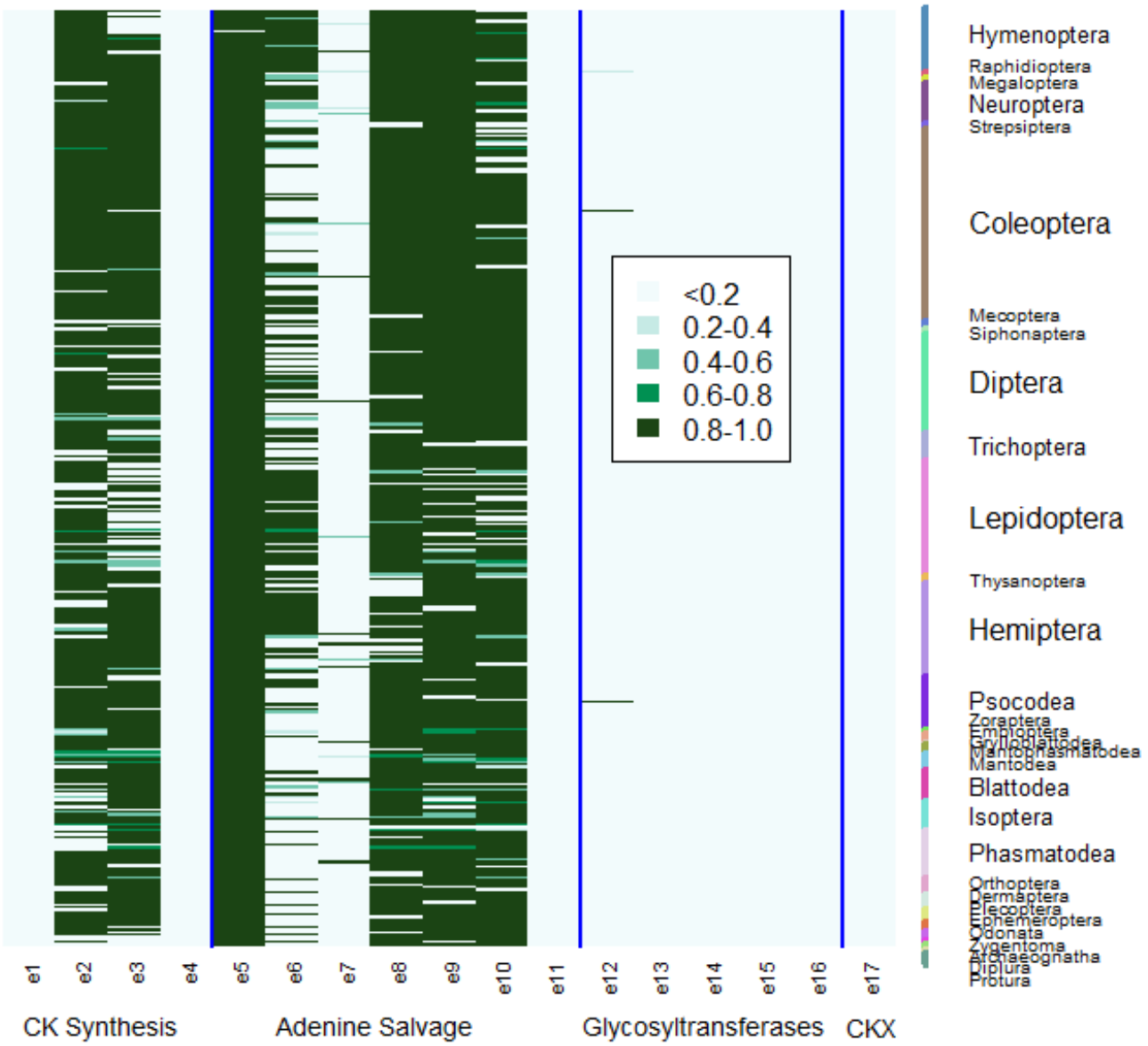
hexapods, but that de novo synthesis, glucosyltransferases, and CK Oxidase genes were not detected in most transcriptomes. See Table 1 for enzyme names and EC numbers.


tRNA
EC:2.5.1.75
tRNA dimethylallyltransferase
DMAPP
A37
i6A37
iPNT
isopentenylation
tRNA degradation
methylthiolation
EC: 2.8.4.3
tRNA-2-methylthio-N6-dimethylallyladenosine
synthase
Thiol
2MeSiPNT
ms2i6A37
S-Adenosyl L-Methionine
tRNA degradation


AMP
EC: 3.2.2.4
e11, AMP nucleosidae
EC: 3.1.3.5
e5, 5'-nucleotidase
EC: 2.7.1.20
e9, adenosine kinase
EC: 2.4.2.7
e10, adenine phosphoribosyltransferase
Adenosine
EC: 3.2.2.1
EC: 3.2.2.7
e6, purine nucleosidase
EC: 2.4.2.1
e8, purine-nucleoside phyosphorylase
EC: 3.2.2.-
e7, cytokinin riboside 5'-monophosphate phosphoribohydrolase
Adenine


Sequence Similarity
KEGG Exemplar Blast
Tentative CK and adenine salvage transcripts
Step 1
Sequence Similarity
Annotated Genomes Blast
Annotations
Filter on Protein Hit Annotations for CK and adenine salvage
Step 2
CK and adenine salvage transcripts
Annotated
Un-characterized Small number
Sequence Similarity & Annotation
Potential Hidden CK transcripts
Blast Proteins NR
Additional CK and adenine salvage transcripts


Kite Transcriptomes
NCBI-Annotated Genomes
All Insect Genomes
Number of Species
0
25
50
75
100
125
150
Coleoptera
Hymenoptera
Neuroptera
Lepidoptera
Diptera
Orthoptera
Hemiptera
Blattodea
Phasmatodea
Psocodea
Raphidioptera
Mantodea
Trichoptera
Mantophasmatodea
Megaloptera
Plecoptera
Dermaptera
Grylloblattodea
Collembola
Embioptera
Ephemeroptera
Isoptera
Odonata
Thysanoptera
Mecoptera
Strepsiptera
Zygentoma
Archaeognatha
Diplura
Psocoptera
Siphonaptera
Protura
Zoraptera


<0.2
0.2-0.4
0.4-0.6
0.6-0.8
0.8-1.0
e1
e2
e3
e4
e5
e6
e7
e8
e9
e10
e11
e12
e13
e14
e15
e16
e17
CK Synthesis
Adenine Salvage
Glycosyltransferases
CKX
Hymenoptera
Raphidioptera
Megaloptera
Neuroptera
Strepsiptera
Coleoptera
Mecoptera
Siphonaptera
Diptera
Trichoptera
Lepidoptera
Thysanoptera
Hemiptera
Psocodea
Zoraptera
Embioptera
Grylloblattodea
Mantophasmatodea
Mantodea
Blattodea
Isoptera
Phasmatodea
Orthoptera
Dermaptera
Plecoptera
Ephemeroptera
Odonata
Zygentoma
Archaeognatha
Diplura
Protura