8-2022

Examination of Changes in Gene Family Size in Carnivorous Plants Through Genomic Analysis

Moonia Ammari

Abstract
The traits associated with carnivory in plants have independently evolved within several orders of Angiosperm. Despite this lack of a common origin, carnivorous plants show many remarkably similar adaptations, suggesting that common evolutionary pressures may have shaped these species. Here, we hypothesize that one common trait carnivorous plants share is the number of genes associated with specific biological functions, such as nutrient transport. To test this hypothesis, we first identified genes and gene families in the model plant species *Arabidopsis thaliana* that are associated with previously characterized biological functions that may be important for plant carnivory. We then compared the number of these genes present in the genomes of carnivorous and non-carnivorous plants, respectively. Our results show that the aquaporin gene family is lost or reduced in carnivorous plants. Additionally, the results also indicate that gene families linked with cysteine-type peptidase, superoxide dismutase, phospholipase, and aspartic-type endopeptidase activity have contracted in size within carnivorous plants relative to non-carnivorous plants. These findings support our hypothesis that the copy number for genes involved in nutrient utilization has diverged between non-carnivorous and carnivorous plants, possibly due to adaptations to similar environmental circumstances in the latter group.

Keywords: Carnivorous plants, convergent evolution, Arabidopsis thaliana, genes, aquaporin, Methylammonium transmembrane transporter activity, water channel activity, TIPs

MONTCLAIR STATE UNIVERSITY

Examination of Changes in Gene Family Size in Carnivorous Plants through Genomic Analysis

By

Moonia Ammari

Master's Thesis Submitted to Faculty of

Montclair State University

In Partial Fulfillment of Requirements

For the Degree of

Master of Science

August 2022

College of Science and Mathematics

Department of Biology

Thesis Committee:

[redacted]

Thesis Sponsor: Dr. Matthew A.a

[redacted]

Committee Member: Dr. James Campanella

[redacted]

Committee Member Dr. Robert Meredith

EXAMINATION OF CHANGES IN GENE FAMILY SIZE IN CARNIVOROUS PLANTS THROUGH GENOMIC ANALYSIS

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

by

Moonia Ammari

Montclair State University

Montclair, NJ

2022

Acknowledgments

First and foremost, I want to thank my thesis adviser, Dr. Matthew L. Aardema, for his inspiration, advice, and support during my whole thesis process. My contribution to science would not have been possible without Dr. Aardema's steadfast counsel and belief in my study. I'd also like to thank my thesis committee members, Dr. James Campanella and Dr. Robert Meredith, for their help, effort, collaboration, and direction. Special appreciation to Montclair State University's Biology department and teachers for their encouragement, advice, and space for my study.

Table of Contents

List of Tables

List of Figures

INTRODUCTION

Convergence in evolution occurs when unrelated species independently acquire identical characteristics, or analogous traits, in response to comparable environmental pressures (Doolittle, 1994; Stern, 2013). Natural selection can independently generate similar morphological and physiological solutions in different taxa (Stern, 2013). This phenomenon, known as "homoplasy", is the consequence of convergent evolution, whereby different species acquire biological structures and functions that appear identical to one another but do not share a common ancestry (Wake, 1991; Stayton, 2015). Convergent evolution is evident in carnivorous plant species, which have independently evolved methods of obtaining nutrients via prey capture rather than directly through their roots (Ellison et al., 2005).

Unique Adaptations of Carnivorous Plants

Carnivorous plants are a diverse, polyphyletic group that are limited to nutrient-deficient habitats (Albert et al., 1992; Hatcher et al., 2020). The 600+ species evolved within five distinct Angiosperm orders: Poales, Caryophyllales, Oxalidales, Ericales, and Lamiales (Givnish et al., 2015; Ellison et al., 2019), with the majority of these being in the Caryophyllales and Lamiales.

Plant carnivory is defined by an ability to detain insects and other organisms in the Kingdom Animalia, metabolically deconstructing and utilizing the resulting products (Hatcher et al., 2020). However, the unique mechanisms used to capture their prey are only one example of the morphological and physiological similarities that exist across carnivorous plants (Biswal et al., 2018). Another example of convergence in these plants occurs is their root systems. In most plant species, roots provide anchoring, water intake, nutritional uptake, and nutrient retention. On the other hand, most carnivorous plants are either hygrophytic (live in water or very moist

soils) or epiphytic (growing on or around other plants), and generally lack a strong structural root network (Adlassnig et al., 2005). Root systems are severely reduced in carnivorous plants like the Venus flytrap (*Dionaea muscipula)*, making them poor competitors for soil resources (Brewer et al., 2003). In some cases, the roots of aquatic carnivorous plants such as the waterwheel (*Aldrovanda vesiculosa)*, have entirely disappeared (Hedrich et al., 2021). In carnivorous plants, the stems and leaves take over the roles of nutrient acquisition, processing, and transport that were formerly performed by the roots (Adlassnig et al., 2005).

Mechanisms of Prey-processing and Nutrient Transfer

Carnivorous plants are often found in moist and acidic soils (Adamec et al., 1997). Such environments result in reduced organic decomposition rates by microorganisms, limiting the availability of nutrients (Rice, 2002). Nitrogen, phosphorus, and potassium are particularly scarce in these environments and this scarcity may select for carnivorous traits (Mithöfer, 2022; Ellison and Gotelli, 2001). Carnivory enables plants to absorb nutrients (directly or indirectly) from prey captured by traps (Adamec, 2013). The comparatively high levels of nitrogen and phosphorus that are present in prey tissue may be processed and utilized for growth and development in carnivorous plants (Adamec, 2008). For example, carnivorous plants can acquire 10% to 80% of their total nitrogen from insects alone (Behie et al., 2013).

The ability of carnivorous plants to break down and digest their prey is a distinct characteristic that enables these plants to obtain nutrients. Specialized enzymes are required for the digestion of prey that promote the absorption of nutrients upon capture. Genes that encode for digestive enzymes are not unique, but rather are closely aligned with ubiquitous gene families found across Angiosperms (Freund et al., 2022). These digestive enzymes are likely an

exaptation of pathogenesis-related proteins involved in plant defense against disease (Ravee et al., 2018). Exaptation is described as an adaptive characteristic that fulfills a purpose other than for what it originated (Gould and Vrba, 1982). Pathogenesis-related genes are frequently secreted proteins that impede fungal and bacterial growth, or function in lipid transfer and defense signaling mechanisms (Freund et al., 2022). Genes that were formerly used by plants to prevent bacterial or fungal infections may now be used to break down or digest prey in carnivorous plants. Such genes may have been co-opted to digest animal material for nutrient consumption as opposed to defense against pests. This route to carnivory may have been especially viable when such genes existed in multiple copies.

Defense-related genes appear to have been repurposed for carnivorous digestive functions to arise. Among the gene families that enable functional digestion are peroxidase, serine-type carboxypeptidase, aspartic proteases, superoxide dismutase, chitinases, proteases, and lipases (Schulze et al., 2012). Peroxidase activity mechanisms break down compounds by utilizing H_2O_2 (Fürstenberg-Hägg et al., 2013). Furthermore, serine-type carboxypeptidase activity works to catalyze functions of defense against herbivores and UV protection (Fraser et al., 2005). Serine-type carboxypeptidase activity in carnivorous plants is associated with the digestion of prey by degrading proteins (Ravee et al., 2018). Aspartic proteases function to ferment and digest in acidic conditions found within the carnivorous plant traps (Kadek et al., 2014). Superoxide dismutase activity is the first line of defense against oxidative stress in plants (Das and Roychoudhury, 2014). Many modified enzymes like chitinase originated as antifungal proteases and lipases to aid in energy storage (Wheeler and Carstens, 2018). Other genes associated with prey digestion are endopeptidase inhibitors (Palfalvi et al., 2020).

In addition to prey digestion, the evolution of carnivory in plants also required novel strategies for nutrient transport. Changes in the expression of genes that code for nutrient transporters are fundamental to the ability of carnivorous plants to get nutrients from their modified leaf traps (Palfalvi et al., 2020). Nutrient transport genes are often found in the roots of normal plants, which are the major location through which the plant gets nutrients directly from the soil. Interestingly, in the digestive glands of carnivorous plants, the genes that encode for nutrient transporters are expressed only after prey has been caught, while expression in non-carnivorous plant root tissues occurs more regularly (Palfalvi et al., 2020). In non-carnivorous plants, nitrogen enters the plant's roots, and aquaporin is a key transporter in water and nitrogen metabolism (Scherzer et al., 2013). Aquaporin belongs to the Major Intrinsic Protein superfamily which are membrane channel proteins that function to transport nutrients like nitrogen, and water for plants to utilize (Afzal et al., 2016). Of the aquaporin subfamily, TIPs (Tonoplast Intrinsic Proteins) transport and store molecules of water, hydrogen peroxide, and glycerol (Regon et al., 2014). The permeability of ammonia or urea to tonoplast aquaporins helps TIPs store and remobilize nitrogen in the vacuoles (Maurel et al., 2015). Genes encoding amino acid transporters have been described in plants, such as *Arabidopsis thaliana* (Fischer et al., 1998). The use of knockout alleles has confirmed that in aquaporins, ammonium membrane transport proteins provide the major route of high-affinity ammonium influx in *A. thaliana* roots (Ludewig et al., 2007). Nutrient uptake is likely to depend on the presence or regulation of specific transporters. The principal source of nitrogen in typical plants is the uptake of charged ammonium ions (NH_4+) (Loqué et al., 2004). Aquaporins are critical to water and nutrient homeostasis and play an important role in plant defenses against both abiotic and biotic systems of plant stress (Afzal et al., 2016). Thus, the flexibility of changes in root hydraulic conductivity

has been linked to the existence and function of aquaporins, which allow water to move across cellular membranes in response to osmotic or hydrostatic pressure gradients (Wang et al. 2019).

In addition to aquaporins, the water channel transporters are transmembrane proteins that form channels to facilitate the rapid movement of water and other small metabolites in either direction of the plasma membrane (Kapilan et al., 2018). The aquaporin's gating behavior is poorly understood, although there is accumulating evidence that phosphorylation, pH, and osmotic gradients may influence water channel function (Volkov et al., 2008). The abundance, or expansion, of the aquaporin gene family and water channel activity, is correlated to water flow across membranes through gene duplication events (Kapilan et al., 2018).

The Evolution of Gene Copy Number

One major route for organisms to evolve novel life history strategies is via the duplication and divergence of specific genes (Zhen et al., 2012). A gene family is several evolutionarily related genes that typically have similar biochemical functions (Demuth and Hahn, 2009; Hao et al., 2019; Ribas de Pouplana, 2020). Many genes may be classified into gene families based on their sequence similarities (Guo, 2013). In most instances, gene duplication of a single original gene results in the creation and expansion of a gene family. Alternatively, genes may be functionally inactivated and eventually lost, resulting in the contraction of a gene family (Hahn et al, 2005). Genes that are important for organismal functioning should be retained whereas genes that have lost their adaptive value may be lost (Wheeler and Carstens, 2018).

Genes important for plant carnivory may be those related to the ability to sense, digest, and uptake nutrients from prey (Hedrich et al., 2018). Such genes may be represented in carnivorous plant genomes in multiple copies. Conversely, due to their reduced or absent

reliance on nutrient uptake from the soil, the number of genes important for nitrogen, potassium, calcium, iron, H_2O transporter, or membrane channel proteins may be reduced (Morgan and Connolly, 2013).

The expansion or contraction of gene families associated with nutrient processing may be a direct product of environmental stress. Examples of gene family expansion in carnivorous plants are an increased number of genes associated with alternative oxidase and ATP: ADP anti-transporter to protect from oxidative stress as opposed to non-carnivorous plants (Wheeler and Carstens, 2018). Duplication events have additionally occurred in gene families associated with trap-specific cysteine proteases, which are critical in carnivorous plant digestion (Lan et al., 2017). Contraction of gene families has been observed in the ubiquitin (UBQ) gene family, which is frequently engaged in stress responses, as well as genes involved in root growth (Palfalvi et al., 2020). Such losses may have adaptive value if the expression of such genes is energetically costly (Aardema et al., 2020).

Hypotheses and Scope

Observations of gene duplication or loss suggest that such genes were either important for the evolution of carnivory or had reduced importance with the emergence of carnivory. By comparing the number of genes in functional categories previously linked to carnivory, it may be feasible to learn more about the general patterns of evolution that led to plant carnivory. Because carnivorous plants developed in nutrient-deficient soils, it is likely that there is a relationship between nutrient utilization and the environmental conditions for carnivory to emerge. Here, we hypothesized that the aquaporin gene family will contract in the genomes of carnivorous plants due to a reduced ability/need to uptake nutrients from the substrate. We further postulated that

additional gene families involved in nutrient uptake of sodium ion transmembrane transporter activity, water channel activity, methyl ammonium transport, symplast, and ammonium transport will exhibit contraction in gene family size, relative to non-carnivorous plants because of new mechanisms of obtaining nutrients through prey. Finally, we predicted that genes involved in digestion will be present in greater numbers (gene family expansion) in carnivorous plants compared to non-carnivorous plants. To test these hypotheses, we employed gene annotations from *A. thaliana,* the best functionally studied plant species, to evaluate potential orthologues based on the number of genes found in the genomes of both carnivorous and non-carnivorous plant species. Annotated genes from *A. thaliana* may enable a better understanding of the functional genes and the mechanisms of contraction or expansion that are present in carnivorous plants.

MATERIALS AND METHODS

Identification of Carnivorous Datasets

We used all publicly available carnivorous plant genome assemblies from the National Center for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov) and the Biozentrum database (https://www.uni-wuerzburg.de/startseite/). *Arabidopsis thaliana* was chosen as the reference species for comparison against all taxa due to its high-quality genome assembly and annotation, as well as widespread knowledge of gene function in this species. The reference protein and transcript sequences were obtained from the website: www.arabidopsis.org. To identify phylogenetically related and well-annotated non-carnivorous control groups, we used the National Center for Biotechnology Information's Taxonomy Browser (Fig.1) (Schoch et al., 2020). The NCBI database was used to gather non-carnivorous taxon assemblies and sequences. We downloaded gene sequence data (amino acid sequences

annotated from assembled genomes) for seven carnivorous taxa and sixteen non-carnivorous taxa (Table 1).

Gene Ontology (GO) designations or codes offer a framework for characterizing functional genes and gene products across all species (Ashburner et al., 2000). Wheeler and Carstens (2018) identified GO terms potentially associated with carnivory (Table 2). Six out of 40 previously determined GO terms were not used in *A. thaliana* annotation. For these GO codes, we located synonymous GO terms using the AmiGO2 Gene Ontology Database (http://www.geneontology.org). The replacements made are detailed in the GO codes and GO term adjustments (Table 3). Next, we identified all genes in *A. thaliana* that were enriched for these GO terms using the file ATH_GO_GOSLIM.txt (Table 4). Genes could be enriched for more than one GO term. As we were predominantly concerned with comparing the relative number of genes associated with specific GO terms between carnivorous and non-carnivorous species, we did not control for differences in the number of genes enriched between GO terms. The amino acid sequences for each gene became our queries for the comparisons between our focal species.

Data Processing and Analysis

With the amino acid sequences from *A. thaliana*, we conducted local protein-to-protein comparisons ('blastp') for each species independently using the program BLAST v. 2.11 (Altschul et al., 1990), with E-values set to 1e-10. With the resulting blast outputs, we determined the number of potential orthologues associated with each of the 40 GO terms (Table 2) using a custom Perl script (Better_Blast2GO.pl). For each species, two separate filtering criteria for identifying likely orthologues associated with each GO term were applied. The first

criteria, designated as “lenient”, required accepted orthologous sequences to have at least 65% sequence similarity with the *A. thaliana* sequence and a match length of 100 amino acids. The second criterion, designated as “conservative”, required protein matches to have at least 75% sequence similarity and a length of 200 amino acids. For each species and GO term combination, the number of unique blast hits matching these criteria were totaled. Gene hits represent putative orthologous sequences associated with a specific GO term and may represent the number of genes found in a functional gene family. Although this did not allow for comparisons between GO terms (as there were different numbers of *A. thaliana* genes enriched for each term), it did allow us to compare the relative number of unique genes within a term present in the carnivorous and non-carnivorous genomes, respectively.

With this data, we examined the relative genomic-level expansion or contraction of genes associated with specific GO terms that were previously identified as potentially important for a carnivorous lifestyle. Genes and gene families that are important for plant carnivory may be expanded in carnivorous taxa relative to non-carnivorous taxa. Conversely, genes and gene families that have reduced importance in carnivorous plants may be lost or reduced. By comparing the number of blast hits matching the above-described criteria for each of the 40 GO terms, it was possible to examine the relative contraction or expansion of gene groups across the examined species. Figure 2 provides a schematic representation of the analytical methodologies described above.

To test the hypothesis that genes related to specific GO terms within carnivorous plants have contracted or expanded, we performed an independent two-sample t-test to compare the number of distinct blast matches between carnivorous and non-carnivorous species for each *A. thaliana* query sequence. This was accomplished using R-studio (Version 1.3.1073) (RStudio,

2015). We recorded the p-value, degrees of freedom, t-value, 95% confidence intervals, and mean values for both carnivorous and non-carnivorous groups in a table for each of the primary gene family categories. For the independent two-sample t-test, we evaluated the lenient and conservative filtering criteria separately.

In statistical testing processes, unequal sample sizes can increase variation between two sample subjects and result in a loss of statistical power (Lopes et al., 2021). To partially account for the unequal number of samples in the carnivorous (n=7) and non-carnivorous groups (n=16), we employed a random subsampling method with replacement to generate datasets with a numerically equivalent number of samples. With these, we performed a two-sample t-test to compare the mean number of blast hits between carnivorous and non-carnivorous groups for each GO term. This subsampling was repeated 100 times for both the lenient and conservative datasets independently using a custom Perl script to randomly subsample the non-carnivorous group each time ('subsample_t_test.pl'; Aardema, unpublished). From the 100 statistical replicates, the number of times that the difference between the carnivorous and non-carnivorous groups were statistically significant (at $p<0.05$) was recorded. This approach allowed us to compare the number of blast hits for GO terms identified in carnivorous and non-carnivorous groups respectively. In the final dataset table, GO term function, p-value, and subsampling replicate counts, for both lenient and conservative criteria were recorded.

Comparative phylogenetic approaches investigate the connections between quantitative features, like gene family expansion or contraction, while accounting for the evolutionary histories of the focal taxa (Tolkoff et al., 2018). Specifically, the gene content of closely related species has a stronger tendency to be similar as a result of common evolution than randomly selected species (Touchman, 2010; Burns and Strauss, 2011). To account for differences in the

evolutionary distance between *A. thaliana* and the taxa examined here (Fig. 1), we performed a third series of analyses, excluding species from the Angiosperm orders Oxidales (*Cephalotus follicularis*), Malpighiales (*Populus trichocarpa*) and Fabales (*Glycine soja*). These species are in the subclass Rosidae (order Brassicales), as is *A. thaliana*. After species removal, we again performed independent t-tests, treating the lenient and conservative datasets separately.

RESULTS

The number of unique blast hits for each taxon and GO term combination are presented in Table 5 for the lenient filtering criteria and Table 6 for the conservative filtering criteria. Two-sample t-tests were used to examine potential differences between carnivorous and non-carnivorous plants for the number of potential orthologous genes associated with specific biological functions (represented by GO terms). A non-significant result in these tests indicated no statistically significant difference in the number of genes associated with a specific function between carnivorous and non-carnivorous plants. Conversely, a statistically different number of blast hits associated with a specific GO term supported the hypothesis that carnivorous plants have undergone changes (expansion or contraction) in these genes.

The filtered number of blastp hits for the aquaporin gene family (GO code 24) in both the lenient ($\bar{x}$=3.09, s=2.98) and conservative ($\bar{x}$=5.30, s=4.36) datasets revealed that the carnivorous species have substantially fewer or no hits (Tables 5 and 6). The number of aquaporin hits from non-carnivorous taxa remained unchanged compared to the carnivorous taxa which had no gene hits, with none indicating any form of loss (Tables 5 and 6). Water channel activity in the conservative analysis revealed gene hits that exhibit considerable reduction in carnivorous plants compared to non-carnivorous plants ($\bar{x}$=127.74, s=80.47). In the conservative

criteria data, a reduced number of gene matches of carnivorous plants present in the aspartic-type endopeptidase activity (x^{-}=11.83, s=15.70), phospholipase activity (x^{-}=3.30, s=2.58) and superoxide dismutase activity (x^{-}=3, s=3). In the lenient criteria data, a reduced number of gene matches of carnivorous plants are present in the aspartic-type endopeptidase (x^{-}=42.35, s=30.54), phospholipase (x^{-}=20.70, s=15.62), and superoxide dismutase activity (x^{-}=40.22, s=26.23).

The subsampling analysis employed 100 sample replicates to assess the statistical significance of the resulting gene matches with replacement. In carnivorous species, the average number of aquaporins (GO code 24) appeared to be reduced in carnivorous taxa relative to non-carnivorous taxa (Table 6). Larger replicate counts of >80 genes showed increased statistical significance between the samples and suggested means of genetic differences between carnivorous and non-carnivorous plants. Aquaporin (counts=88) and water channel activity (counts=82) supported the observation that replicates with more than 80 counts showed statistical significance in the conservative datasets (Fig. 3). In the lenient datasets, the GO terms that had greater than 80 replicate counts were aquaporin (counts=96), serine-type carboxypeptidase (counts=89), peroxidase (counts=95), and phospholipase activity (counts=89). The GO terms in the conservative dataset with >40 counts deemed to have some statistical significances were, aspartic-type endopeptidase (counts=49), phospholipase (counts= 45), superoxide dismutase (counts=45), and cysteine-type peptidase activity (counts= 43). The GO terms in the lenient dataset with >40 counts were superoxide dismutase (counts=63), lipid transport (counts=48), ammonium transmembrane transport (counts= 44), alternative oxidase (counts=47), ATPase (counts= 40), heat shock protein (counts=61), protein homodimerization (counts=60), cinnamyl-alcohol dehydrogenase (counts=43), Glucan endo-1,3-beta-glucanase (counts=43), and ammonium membrane transport (counts= 54). All other GO terms with a replicate count of <40

were considered to have no statistical significance between carnivorous and non-carnivorous plants and indicated that there are no biological differences between the two.

The results for the comparative phylogenetic control analysis are presented in Tables 7 and 8 for the lenient and conservative criteria respectively. The number of gene hits for aquaporin remained decreased or lost in the lenient analysis, comparable to when the taxa were present in the previous analysis ($\bar{x}$=4.90, s=4.36). The number of genes were reduced in comparison to the non-carnivorous taxa for water channel ($\bar{x}$=278.60 s=130.98), phospholipase, ($\bar{x}$=17.30, s=10.18), and superoxide activity ($\bar{x}$=35.3 s=20.84). In the conservative analysis, the number of genes were additionally lost or reduced compared to the non-carnivorous taxa ($\bar{x}$=2.60, s=2.82). Furthermore, water channel continued to exhibit signs of reduced size ($\bar{x}$=116.5, s= 72.5), along with aspartic-type endopeptidase ($\bar{x}$= 7.85, s=7.58), phospholipase ($\bar{x}$=2.95, s=2.39) and superoxide dismutase activity ($\bar{x}$=2.90, s= 3.08).

DISCUSSION

Carnivorous plants appeared to have a lower number of copies for genes associated with nutrient uptake and digestion compared to non-carnivorous plants. In particular, the aquaporin gene family appeared to have contracted significantly, based on our analyses with both the lenient and conservative datasets (Tables 5 and 6). In addition, while genes related to water channel activity remain present in carnivorous plants, there was evidence for gene copy reduction. Genes related to digestion such as aspartic-type endopeptidases, phospholipases, peroxidases, cysteine-type peptidases, and superoxide dismutases also showed indications of contraction. Other nutrient transporters like sodium ion transmembrane transporter, methyl ammonium transport, symplast, and ammonium transport did not show signs of contraction. The

contraction found in the number of genes related to nutrient intake revealed biological differences among carnivorous and non-carnivorous plants.

Aquaporins and other water transporters

Aquaporins mediate water transfer and utilization of nutrients like nitrogen (Forrest and Bhave, 2007). For example, types of aquaporins (like PIP1) may operate as transporters for tiny solutes/gasses, or they may require activation in the plant to function as water channels (Kaldenhoff and Fischer, 2006). Notably, the aquaporin gene family has contracted in carnivorous plants, which may be an adaptation to the reduced supply of nutrients found in the carnivorous plant environments. Most nutrient transporters and channel activities are found in plant roots; however, since carnivorous plants have significantly reduced or missing root systems, selection to maintain the functioning of these transporters may be relaxed, resulting in the inactivation and loss of aquaporin genes (Vatansever et al., 2017). The genes of *A. thaliana* that encode for aquaporin function include methylammonium transmembrane transporter activity (TIP2-3), whose major role is to carry nitrogen supply from the root to be used. These results suggest that carnivorous plants have undergone this loss owing to a lack of root systems and a necessity for nutrient absorption through roots. This is suggested because this transporter displays a reduced number of copies in carnivorous plants. Specific ecological conditions that resulted in these genes no longer being necessary for organismal fitness may explain aquaporin losses within carnivorous plants (Demuth and Hahn, 2009). However, several abiotic variables, including salt, low temperatures, wounding, and drought, also have been shown to influence the abundance of genes in the aquaporin gene family (Kapilan et al., 2018).

The number of genes for water channel activity, although remaining present, show signs of contraction in comparison to non-carnivorous plants. The differences in abundance for these genes are likely due to alterations in the nutrient utilization of the roots. To optimize access to water, the root system adapts to the pattern of water availability (Lind et al., 2021). Selection for the loss of some of the water channel transport genes is most likely caused by reduced or lost root systems in carnivorous plants. Aquaporins and water channel transport may be linked functionally, which may be why the water channel activity gene families show biological variations in gene counts when compared to non-carnivorous taxa. This suggests that any changes in the aquaporin gene family may be related to the importance of water channel function. A low abundance of aquaporins in soil stress, like limited nutrients, can reduce available water transport rates (Kapilan et al., 2018). The contraction of aquaporin-related gene families may have negotiated some form of selection response for water channel activity as a facilitated mode of nutrient and water transport.

Other Gene families Suggesting Evolutionary Mechanisms at Play

The results of this study indicate that the number of digestive enzymes, specifically cysteine-type peptidases, superoxide dismutase, phospholipases, and aspartic-type endopeptidases are reduced in carnivorous plants compared to non-carnivorous plants (Table 5). This result contradicts our hypothesis that carnivorous plant genes involved in digestion would expand. It is possible that some carnivorous species have evolved novel compounds for the purposes of prey digestion (Hatcher et al., 2020), as opposed to the co-option of genes important in plant defense. Combined with this novel gene evolution, carnivorous plants may have fewer

numbers of these genes due to a reduced pathogen burden in the habitats where they thrive (Miguel et al., 2019).

Limitations of the study

In our genomic analysis, we used publicly available genomic sequences of both carnivorous and non-carnivorous species. As additional sequencing data becomes available, one route to explore would be to include genomic sequences of new carnivorous plants. The inclusion of additional carnivorous plant taxa could allow for a more accurate understanding of patterns in gene family expansion or contraction. Additionally, more sequences of carnivorous plant data would allow a deeper understanding of how gene contraction or expansion events relate to carnivory.

Furthermore, the implication of additional plant species that are in the same nutrient-poor environments can provide insight as to how selection can be favored for the carnivorous trait. The plants that do persist in the same environment as carnivorous plants suggest that adaptations for this lifestyle are not necessarily the only way to survive. There might be a selection of other gene families that do allow them to persist in such environments. Utilizing additional non-carnivorous sequencing data would allow pinpointing how and why selection may lead to the carnivorous phenotype whereas others would not. This could accurately assess the idea that environmental stresses could lead to alterations or biological differences in genetics to survive and adapt through time. Moreover, it would identify which genes or gene families can potentially lead to the structure and function of the trapping mechanisms.

CONCLUSION

In conclusion, when utilizing *A. thaliana* genes as a reference, gene families of carnivorous plants appear to have contracted in size when compared to other non-carnivorous plant taxa. One of the conditions that have enabled natural selection to favor carnivorous plants is a shortage of nutrients in the soil, which has resulted in the deletion of several genes essential for nutrient digestion and absorption. Overall, our hypothesis is supported by the substantial contraction of the aquaporin gene family. Other gene families involved in nutrient transport, such as water channel activity, did reduce, although not considerably. Gene families that code for proteolytic digestive enzymes, cysteine peptidases, superoxide dismutase, phospholipases, and aspartate endopeptidases, contracted as well in comparison to typical plants. Furthermore, we discovered that additional gene families involved in nutrient uptakes, such as sodium ion transmembrane transporter activity, methyl ammonium transport, symplast, and ammonium transport, did not show contraction or expansion in gene family size when compared to non-carnivorous plants. Data mining and genomic analysis are useful ways to compare the number of genes in plants to determine whether functional groupings or gene families have contracted or expanded enabling carnivory in plants to occur.

REFERENCES

Aardema, M. L., Stiassny, M. L., & Alter, S. E. (2020). Genomic analysis of the only blind cichlid reveals extensive inactivation in eye and pigment formation genes. *Genome biology and evolution*, *12*(8); 1392-1406.

Adamec, L. (2002). Leaf absorption of mineral nutrients in carnivorous plants stimulates root nutrient uptake. *New Phytologist*, *155*(1); 89-100.

Adamec, L. (2008). Mineral nutrient relations in the aquatic carnivorous plant Utricularia australis and its investment in carnivory. *Fundamental and Applied Limnology*, *171*(3); 175-183.

Adamec, L. (1997) Mineral nutrition of carnivorous plants: A review. *Bot. Rev,* (63); 273–299

Adlassnig, W., Peroutka, M., Lambers, H., & Lichtscheidl, I. K. (2005). The roots of carnivorous plants. *Plant and Soil*, *274*(1); 127-140.

Adlassnig, W., Steinhauser, G., Peroutka, M., Musilek, A., Sterba, J. H., Lichtscheidl, I. K., & Bichler, M. (2009). Expanding the menu for carnivorous plants: uptake of potassium, iron, and manganese by carnivorous pitcher plants. *Applied Radiation and Isotopes*, *67*(12); 2117-2122.

Afzal, Z., Howton, T. C., Sun, Y., & Mukhtar, M. S. (2016). The Roles of Aquaporins in Plant Stress Responses. *Journal of developmental biology*, *4*(1); 1-22.

Albert, V. A., Williams, S. E., & Chase, M. W. (1992). Carnivorous plants: phylogeny and structural evolution. *Science*, 257(5076); 1491-1495.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3); 403–410.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: a tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, *25*(1); 25–29.

Behie, S. W., & Bidochka, M. J. (2013). Insects as a Nitrogen Source for Plants. *Insects*, 4(3); 413–424.

Biswal, D. K., Debnath, M., Konhar, R., Yanthan, S., & Tandon, P. (2018). Phylogeny and biogeography of carnivorous plant family Nepenthaceae with reference to the Indian pitcher plant Nepenthes khasiana reveals an Indian subcontinent origin of Nepenthes colonization in Southeast Asia during the Miocene epoch. *Frontiers in Ecology and Evolution*, 6(180); 1-12.

Brewer, J. S. (2003). Why don't carnivorous pitcher plants compete with non-carnivorous plants for nutrients. *Ecology*, *84*(2); 451-462.

Burns, J. H., & Strauss, S. Y. (2011). More closely related species are more ecologically similar in an experimental test. *Proceedings of the National Academy of Sciences*, *108*(13); 5302-5307.

Chang & Gallie (1997) Chang SC, Gallie DR. *(1997)* RNase activity decreases following a heat shock in wheat leaves and correlates with its posttranslational modification. *Plant Physiology*. 113; 1253–1263.

Cheng, G., Wang, M., Zhang, L., Wei, H., Wang, H., Lu, J., & Yu, S. (2022). Overexpression of a Cotton Aquaporin Gene GhTIP1; 1-like Confers Cold Tolerance in Transgenic Arabidopsis. *International Journal of Molecular Sciences,* 23(3), 1-20.

Christin, P. A., Weinreich, D. M., & Besnard, G. (2010). Causes and evolutionary significance of genetic convergence. *Trends in Genetics*, *26*(9); 400-405.

Das, K., & Roychoudhury, A. (2014). Reactive oxygen species (ROS) and the response of antioxidants as ROS-scavengers during environmental stress in plants. *Frontiers in environmental science,* 2(53); 1-13.

Demuth, J. P., & Hahn, M. W. (2009). The life and death of gene families. *Bioessays*, *31*(1); 29-39.

Doolittle, R. F. (1994). Convergent evolution: the need to be explicit. *Trends in biochemical sciences*, 19(1); 15-18.

Ellison A. M. & Adamec Lubomír. (2018). Carnivorous plants: physiology ecology and evolution (First). *Oxford University Press,* Book, (3); 22-220.

Ellison, A. M., & Gotelli, N. J. (2009). Energetics and the evolution of carnivorous plants—Darwin's 'most wonderful plants in the world'. *Journal of Experimental Botany,* 60(1); 19-42.

Ellison, A. M., & Gotelli, N. J. (2001). Evolutionary ecology of carnivorous plants. *Trends in ecology & evolution*, *16*(11); 623-629.

Ellison, Aaron & Gotelli, Nicholas & Brewer, Stephen & Cochran-Stafira, Liane & Kneitel, Jamie & Miller, Thomas & Worley, Anne & Zamora, Regino (2001). The evolutionary ecology of carnivorous plants. *Advances in Ecological Research*. 33; 1-74.

Estrella, L., & Albert, V. A. (2017). Long-read sequencing uncovers the adaptive topography of a carnivorous plant genome. *Proceedings of the National Academy of Sciences of the United States of America,* 114(22); E4435–E4441.

Fleischmann, A., Schlauer, J., Smith, S. A., Givnish, T. J., Ellison, A. M., & Adamec, L. (2018). Evolution of carnivory in angiosperms. *Carnivorous plants: physiology, ecology, and evolution*, Book; 22-42.

Forrest, K. L., & Bhave, M. (2007). Major intrinsic proteins (MIPs) in plants: a complex gene family with major impacts on plant phenotype. *Functional & integrative genomics*, *7*(4); 263-289.

Fraser, C. M., Rider, L. W., & Chapple, C. (2005). An expression and bioinformatics analysis of the Arabidopsis serine carboxypeptidase-like gene family. *Plant Physiology,* 138(2); 1136–1148.

Freund, M., Graus, D., Fleischmann, A., Gilbert, K. J., Lin, Q., Renner, T., ... & Fukushima, K. (2022). The digestive systems of carnivorous plants. *Plant Physiology*; 1-16.

Fukushima, K., Fang, X., Alvarez-Ponce, D., Cai, H., Carretero-Paulet, L., Chen, C., ... & Hasebe, M. (2017). The genome of the pitcher plant Cephalotus reveals genetic changes associated with carnivory. *Nature Ecology & Evolution*, *1*(3); 1-9.

Fürstenberg-Hägg, J., Zagrobelny, M., & Bak, S. (2013). Plant defense against insect herbivores. *International journal of molecular sciences*, *14*(5); 10242–10297.

Gibson, T. C., & Waller, D. M. (2009). Evolving Darwin's 'most wonderful'plant: ecological steps to a snap-trap. *New Phytologist*, *183*(3); 575-587.

Givnish, T. J., Burkhardt, E. L., Happel, R. E., & Weintraub, J. D. (1984). Carnivory in the bromeliad *Brocchinia reducta*, with a cost/benefit model for the general restriction of carnivorous plants to sunny, moist, nutrient-poor habitats. *The American Naturalist*, *124*(4); 479-497.

Givnish, T. J. (2015). New evidence on the origin of carnivorous plants. *Proceedings of the National Academy of Sciences*, *112*(1); 10-11.

Gould, S., & Vrba, E. (1982). Exaptation—a Missing Term in the Science of Form. *Paleobiology, 8*(1); 4-15.

Hao, Y., Qu, Y., Song, G., & Lei, F. (2019). Genomic Insights into the Adaptive Convergent Evolution. *Current genomics*, *20*(2); 81–89.

Hahn, M. W., De Bie, T., Stajich, J. E., Nguyen, C., & Cristianini, N. (2005). Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genome research*, *15*(8); 1153-1160.

Hedrich, R., & Fukushima, K. (2021). On the origin of carnivory: Molecular physiology and evolution of plants on an animal diet. *Annual Review of Plant Biology*, *72*; 133-153.

Hedrich, R., & Neher, E. (2018). Venus flytrap: how an excitable, carnivorous plant works. *Trends in Plant Science*, *23*(3); 220-234.

Huang, R., O'Donnell, A. J., Barboline, J. J., & Barkman, T. J. (2016). Convergent evolution of caffeine in plants by co-option of exapted ancestral enzymes. *Proceedings of the National Academy of Sciences*, *113*(38); 10613-10618.

Jakšová, J., Adamec, L., Petřík, I., Novák, O., Šebela, M., & Pavlovič, A. (2021). Contrasting effect of prey capture on jasmonate accumulation in two genera of aquatic carnivorous plants (Aldrovanda, Utricularia). *Plant physiology and biochemistry* : PPB, 166; 459–465.

Kadek, A., Tretyachenko, V., Mrazek, H., Ivanova, L., Halada, P., Rey, M., Schriemer, D. C., & Man, P. (2014). Expression and characterization of plant aspartic protease nepenthesin-1 from Nepenthes gracilis. *Protein expression and purification*, 95;121–128.

Kaldenhoff, R., & Fischer, M. J. A. P. (2006). Aquaporins in plants. *Acta Physiologica*, *187*(1-2); 169-176.

Kapilan, R., Vaziri, M., & Zwiazek, J. J. (2018). Regulation of aquaporins in plants under stress. *Biological research*, *51*(1); 1-11

Lan, T., Renner, T., Ibarra-Laclette, E., Farr, K. M., Chang, T. H., Cervantes-Pérez, S. A., Zheng, C., Sankoff, D., Tang, H., Purbojati, R. W., Putra, A., Drautz-Moses, D. I., Schuster, S. C., Herrera, Ludewig, U., Neuhäuser, B., & Dynowski, M. (2007). Molecular mechanisms of ammonium transport and accumulation in plants. *FEBS Letters*, *581*(12); 2301-2308.

Lan, T., Renner, T., Ibarra-Laclette, E., Farr, K. M., Chang, T. H., Cervantes-Pérez, S. A., ... & Albert, V. A. (2017). Long-read sequencing uncovers the adaptive topography of a carnivorous plant genome. *Proceedings of the National Academy of Sciences*, *114*(22); E4435-E4441.

Lind, K. R., Siemianowski, O., Yuan, B., Sizmur, T., VanEvery, H., Banerjee, S., & Cademartiri, L. (2021). Evidence for root adaptation to a spatially discontinuous water availability in the absence of external water potential gradients. *Proceedings of the National Academy of Sciences*, *118*(1); e2012892118.

Lopes, I., Altab, G., Raina, P., & De Magalhães, J. P. (2021). Gene size matters: an analysis of gene length in the human genome. *Frontiers in genetics*, *12(559998)*; 1-15.

Loqué, D., & von Wirén, N. (2004). Regulatory levels for the transport of ammonium in plant roots. *Journal of experimental botany*, *55*(401); 1293-1305.

Maurel, C., Boursiac, Y., Luu, D. T., Santoni, V., Shahzad, Z., & Verdoucq, L. (2015). Aquaporins in plants. *Physiological Reviews*, *95*(4); 1321-1358.

Miguel, S., Nisse, E., Biteau, F., Rottloff, S., Mignard, B., Gontier, E., Hehn, A., & Bourgaud, F. (2019). Assessing Carnivorous Plants for the Production of Recombinant Proteins. *Frontiers in plant science*, 10, (793): 1-13.

Mithöfer, A. (2022). Carnivorous plants and their biotic interactions. *Journal of Plant Interactions*, *17*(1); 333-343.

Morgan, J. Á., & Connolly, E. Á. (2013). Plant-soil interactions: nutrient uptake. *Nature Education Knowledge*, *4*(8); 2-23.

Olson, M. V. (1999). When less is more: gene loss as an engine of evolutionary change. *The American Journal of Human Genetics*, *64*(1); 18-23.

Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R., Moret, B. M., & Stamatakis, A. (2010). How many bootstrap replicates are necessary. *Journal of computational biology*, *17*(3); 185-200.

Pavlovič, A., & Mithöfer, A. (2019). Jasmonate signaling in carnivorous plants: a copycat of plant defense mechanisms. *Journal of experimental botany*, *70*(13); 3379-3389.

Palfalvi, G., Hackl, T., Terhoeven, N., Shibata, T. F., Nishiyama, T., Ankenbrand, M., ... & Hedrich, R. (2020). Genomes of the Venus flytrap and close relatives unveil the roots of plant carnivory. *Current Biology*, *30*(12); 2312-2320.

Płachno, B., Kozieradzka-Kiszkurno, M., Piotr, Ś., & Darnowski, D. W. (2008). Prey attraction in carnivorous Genlisea (Lentibulariaceae). *Acta Biologica Cracoviensia. Series Botanica*, *50*(2); 87-94.

Ravee, R., Mohd Salleh, F. '., & Goh, H. H. (2018). Discovery of digestive enzymes in carnivorous plants with focus on proteases. *PeerJ*, 6; e4914.

Regon, P., Panda, P., Kshetrimayum, E., & Panda, S. K. (2014). Genome-wide comparative analysis of tonoplast intrinsic protein (TIP) genes in plants. *Functional & integrative genomics*, *14*(4); 617–629.

Ribas de Pouplana L. (2020). The evolution of aminoacyl-tRNA synthetases: From dawn to LUCA. *The Enzymes*, *48*; 11–37.

Rice, B. (2002). Carnivorous plants. *Biologist*, 49(6); 245-249.

RStudio Team. (2015). *RStudio: Integrated Development Environment for R*. Boston, MA.

Sackton, T. B., & Clark, N. (2019). Convergent evolution in the genomics era: new insights and directions. *Phil. Trans. R. Soc.* B 374; 20190102.

Scherzer, S., Krol, E., Kreuzer, I., Kruse, J., Karl, F., von Rüden, M., ... & Hedrich, R. (2013). The Dionaea muscipula ammonium channel DmAMT1 provides NH4+ uptake associated with Venus flytrap's prey digestion. *Current Biology*, *23*(17); 1649-1657.

Schulze, W. X., Sanggaard, K. W., Kreuzer, I., Knudsen, A. D., Bemm, F., Thøgersen, I. B., Bräutigam, A., Thomsen, L. R., Schliesky, S., Dyrlund, T. F., Escalante-Perez, M., Becker, D., Schultz, J., Karring, H., Weber, A., Højrup, P., Hedrich, R., & Enghild, J. J. (2012). The protein composition of the digestive fluid from the venus flytrap sheds light on prey digestion mechanisms. *Molecular & cellular proteomics* : MCP, 11(11); 1306–1319.

Schoch, C. L., Ciufo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J. P., Sun, L., Turner, S., & Karsch-Mizrachi, I. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database: the journal of biological databases and curation*, 1; 1-21.

Schlauer, Jan. (2010). Carnivorous Plant Systematics. *Carnivorous Plant Newsletter*. 39; 8-24.

Stayton C. T. (2015). What does convergent evolution mean? The interpretation of convergence and its implications in the search for limits to evolution. *Interface focus*, *5*(6), 20150039.

Stern, D. L. (2013). The genetic causes of convergent evolution. *Nature Reviews Genetics*, 14(11), 751-764.

Touchman, J. (2010). Comparative genomics. *Nature Education Knowledge*, *3*(10), 1-13.

Tolkoff, M. R., Alfaro, M. E., Baele, G., Lemey, P., & Suchard, M. A. (2018). Phylogenetic factor analysis. *Systematic biology*, *67*(3), 384-399.

Veleba, A., Zedek, F., Horová, L., Veselý, P., Srba, M., Šmarda, P., & Bureš, P. (2020). Is the evolution of carnivory connected with genome size reduction. *American Journal of Botany*, *107*(9); 1253-1259.

Wake, D. B. (1991). Homoplasy: The Result of Natural Selection, or Evidence of Design Limitations? *The American Naturalist*, 138(3); 543–567.

Wheeler, G. L., & Carstens, B. C. (2018). Evaluating the adaptive evolutionary convergence of carnivorous plant taxa through functional genomics. *PeerJ*, 6; e4322.

Xu, Y. C., Niu, X. M., Li, X. X., He, W., Chen, J. F., Zou, Y. P., ... & Guo, Y. L. (2019). Adaptation and phenotypic diversification in Arabidopsis through loss-of-function mutations in protein-coding genes. *The Plant Cell*, *31*(5); 1012-1025.

Zamora, R., Gómez, J. M., & Hódar, J. A. (1998). Fitness responses of a carnivorous plant in contrasting ecological scenarios. *Ecology*, *79*(5); 1630-1644.

Zhen, Y., Aardema, M. L., Medina, E. M., Schumer, M., & Andolfatto, P. (2012). Parallel molecular evolution in an herbivore community. *science*, *337*(6102); 1634-1637.

Table 1. The taxonomic classification of Angiosperm species used based on Order.

Order	Family	Genus	Species	Common name	Data found	Designation Codes
Caryophyllales	Droseraceae	*Aldrovanda*	*vesiculosa*	waterwheel	biozentrum.uni-wuerzburg.de	58542
		Dionaea	*muscipula*	Venus flytrap	biozentrum.uni-wuerzburg.de	58543
		Drosera	*spatulata*	spoon-leaved sundew	biozentrum.uni-wuerzburg.de	58541
	Polygonaceae	*Fagopyrum*	*esculentum*	common buckwheat	ncbi.nlm.nih.gov	GCA_001661195.1
	Amaranthaceae	*Beta*	*vulgaris*	beet	ncbi.nlm.nih.gov	GCA_000511025.2
		Chenopodium	*quinoa*	quinoa	ncbi.nlm.nih.gov	GCA_001683475.1
		Spinacia	*oleracea*	spinach	ncbi.nlm.nih.gov	GCA_002007265.1
Lamiales	Lentibulariaceae	*Utricularia*	*gibba*	humped/floating bladderwort	ncbi.nlm.nih.gov	GCA_002189035.1
		Genlisea	*aurea*	corkscrew	ncbi.nlm.nih.gov	GCA_000441915.1
		Utricularia	*reniformis*	bladderwort	ncbi.nlm.nih.gov	GCA_009725065.1
	Gesneriaceae	*Dorcoceras*	*hygrometricum*	queensland rock violet	ncbi.nlm.nih.gov	GCA_001598015.1
	Pedaliaceae	*Sesamum*	*indicum*	sesame	ncbi.nlm.nih.gov	GCA_000512975.1
	Oleaceae	*Olea*	*europaea*	common olive	ncbi.nlm.nih.gov	GCA_002742605.1
	Lamiaceae	*Salvia*	*splendens*	scarlet sage	ncbi.nlm.nih.gov	GCA_004379255.1
	Phrymaceae	*Erythranthe*	*guttata*	spotted monkey flower	ncbi.nlm.nih.gov	GCA_000504015.1
	Bignoniaceae	*Handroanthus*	*impetiginosus*	pink trumpet tree	ncbi.nlm.nih.gov	GCA_002762385.1
	Orobanchaceae	*Striga*	*asiatica*	asiatic witchweed	ncbi.nlm.nih.gov	GCA_008636005.1
		Phtheirospermum	*japonicum*	-	ncbi.nlm.nih.gov	GCA_014905375.1
	Gesneriaceae	*Dorcoceras*	*hygrometricum*	-	ncbi.nlm.nih.gov	GCA_001598015.1
Solanales	Solanaceae	*Solanum*	*lycopersicum*	tomato	ncbi.nlm.nih.gov	GCA_000188115.3
Gentianales	Rubiaceae	*Coffea*	*canephora*	robusta coffee	ncbi.nlm.nih.gov	GCA_900059795.1
Oxalidales	Cephalotaceae	*Cephalotus*	*follicularis*	western australian pitcher	ncbi.nlm.nih.gov	GCA_001972305.1
	Elaeocarpaceae	*Aristotelia*	*chilensis*	chilean wineberry	ncbi.nlm.nih.gov	GCA_008126665.1
Malpighiales	Salicaceae	*Populus*	*trichocarpa*	black cottonwood	ncbi.nlm.nih.gov	GCA_000002775.3
Fabales	Fabaceae	*Glycine*	*soja*	wild soybean	ncbi.nlm.nih.gov	GCA_004193775.2
Brassicaceae	Brassicaceae	*Arabidopsis*	*thaliana*	thale cress	www.arabidopsis.org	GCA_000001735.2

Table 2. List of GO codes and gene-associated functions of carnivorous plants.

Gene Ontology	term GO code	Gene Ontology term	GO code
actin filament	GO:0005884	*heat shock protein activity*	GO:0042026; GO:0006986; GO:0034620
alpha-galactosidase activity	GO:0004557	*lipase activity*	GO:0016298
alternative oxidase activity	GO:0009916	*lipid transport*	GO:0006869
ammonium transmembrane transport	GO:0008519 GO:0072488	*Aquaporin methylammonium transmembrane transporter activity*	GO:0015200
aspartic-type endopeptidase activity	GO:0004190	*peroxidase activity*	GO:0004601
ATP:ADP antiporter activity	GO:0005471	*phosphatase activity*	GO:0016791
ATPase activity	GO:0016887	*phospholipase activity*	GO:0004620
beta-galactosidase activity	GO:0004565	*polygalacturonase activity*	GO:0004650
Glucan endo-1,3-beta-glucanase activity	GO:0052861	*polygalacturonase inhibitor activity*	GO:0090353
chitinase activity	GO:0004568	*protein homodimerization activity*	GO:0042803
cinnamyl-alcohol dehydrogenase activity	GO:0045551	*ribonuclease activity*	GO:0004540
cyclic-nucleotide phosphodiesterase activity	GO:0004112	*serine-type carboxypeptidase activity*	GO:0004185

cysteine-type peptidase activity	GO:0008234	*Sodium ion transmembrane transporter activity*	GO:0015081
endonuclease activity	GO:0004519	*superoxide dismutase activity*	GO:0004784
formate dehydrogenase complex	GO:0009326	*symplast*	GO:0055044
fructose-bisphosphate aldolase activity	GO:0004332	*thioglucosidase activity*	GO:0019137
glucosidase activity	GO:0015926	*water channel activity*	GO:0015250
glutathione transferase activity	GO:0004364	*Endo-1,4-beta-xylanase activity*	GO:0031176

Table 3. GO code and GO term alternatives were found for the reference *Arabidopsis thaliana*.

GO Codes (Wheeler and Carstens, 2018	Alt. GO Codes	GO Term Description
GO:1902687		glucosidase complex
	GO:0015926	glucosidase activity
GO:1905348		endonuclease complex
	GO:0004519	endonuclease activity
GO:0052736		beta-glucanase activity
	GO:0052861	glucan endo-1,3-beta-glucanase activity
GO:0015264		methylammonium channel activity
	GO:0015843	methylammonium transport
	GO:0015200	methylammonium transmembrane transporter activity
GO:0022816		sodium ion transmembrane transporter activity
	GO:0015081	sodium ion transmembrane transporter activity
GO:0097599		xylanase activity
	GO:0031176	endo-1,4-beta-xylanase activity

Table 4. Genes in *A*. thaliana that correspond to the GO codes listed.

GO Codes	Corresponding A. thaliana Genes
GO:0004112	AT4G18930
GO:0004185	AT1G11080, AT2G22960, AT3G10410, AT4G12910 AT1G15000, AT2G22970, AT3G10450, AT4G15100 AT1G28110, AT2G22980, AT3G12203, AT4G30610 AT1G33540, AT2G22990, AT3G12220, AT4G30810, AT1G43780, AT2G23000, AT3G12230, AT5G08260 AT1G61130, AT2G23010, AT3G12240, AT5G09640 AT1G73270, AT2G24000, AT3G17180, AT5G22960 AT1G7328, AT2G24010, AT3G25420, AT5G22980 AT1G73290, AT2G27920, AT3G45010, AT5G23210 AT1G73300, AT2G33530, AT3G52000, AT5G36180 AT1G73310, AT2G35770, AT3G52010, AT5G42230 AT2G05850, AT2G35780, AT3G52020, AT5G42240 AT2G12480, AT3G02110, AT3G56540, AT2G22920, AT3G07990, AT3G63470
GO:0004190	AT1G01300, AT1G79720, AT3G51340, AT5G36260, AT1G01650, AT2G28010, AT3G51360, AT5G48430, AT1G03220, AT2G28030, AT3G54400, AT1G03230, AT2G28040, AT3G61820, AT1G05820, AT2G29900, AT4G30030, AT1G05840, AT2G36670, AT4G30040, AT1G08210, AT2G39710, AT4G33490, AT1G08700, AT2G42980, AT5G02190, AT1G25510, AT2G43070, AT5G07030, AT1G49050, AT3G12700, AT5G19100, AT1G63690, AT3G13235, AT5G19110, AT1G64830, AT3G25700, AT5G19120, AT1G65240, AT3G42550, AT5G22850, AT1G66180, AT3G51330, AT5G33340,
GO:0004332	AT2G01140. AT2G21330, AT2G36460, AT3G52930, AT4G26520, AT4G26530, AT4G38970, AT5G03690
GO:0004364	AT1G02920, AT1G59670, AT2G02390, AT3G43800 AT5G44990, AT1G02930, AT1G59700, AT2G02930, AT3G47680 AT5G45020 AT1G02940, AT1G65820, AT2G29420, AT3G55040 AT5G62480 AT1G02950, AT1G69920, AT2G29440, AT3G62760, AT1G10360, AT1G69930, AT2G29450, AT4G02520, AT1G10370, AT1G74590, AT2G29460, AT4G19880, AT1G17170, AT1G75270, AT2G29470, AT5G02780, AT1G17180, AT1G77290, AT2G29480, AT5G02790, AT1G17190, AT1G78320, AT2G29490, AT5G16710, AT1G19570, AT1G78340, AT2G30860, AT5G17220, AT1G27130, AT1G78360, AT2G30870, AT5G41210, AT1G27140, AT1G78370, AT2G47730, AT5G41220, AT1G49860, AT1G78380, AT3G03190, AT5G41240, AT1G53680, AT2G02380, AT3G09270, AT5G44000,
GO:0004519	AT1G11190, AT4G21590, AT1G11290, AT4G21600, AT1G19100, AT4G24970, AT1G30460, AT4G30870 AT1G59720, AT4G36280, AT1G65070, AT4G36290, AT1G68290, AT5G12220, AT2G15820, AT5G13130 AT2G17510, AT5G50780, AT2G21800, AT5G54090, AT2G22140, AT5G63420, AT2G41460, AT3G28030, AT4G21585,
GO:0004540	AT1G53850, AT2G02990, AT2G04270, AT2G17510, AT2G43190, AT3G04720, AT3G14290, AT4G25630, AT5G24360
GO:0004557	AT3G56310, AT5G08370
GO:0004565	AT1G31740, AT3G53080, AT1G45130, AT3G54440, AT1G72990, AT4G26140, AT1G77410, AT4G35010, AT2G04060, AT4G36360, AT2G04062, AT4G38590, AT2G16730, AT5G20710, AT2G28470, AT5G56870, AT2G32810, AT5G63800, AT3G13750, AT5G63810, AT3G49880, AT5G67540, AT3G52840, AT3G53050, AT3G53075,
GO:0004568	AT1G02360, AT4G01700, AT1G05850, AT4G19720, AT1G56680, AT4G19730, AT2G43570, AT4G19740, AT2G43580, AT4G19750, AT2G43590, AT4G19760, AT2G43600, AT4G19770, AT2G43610, AT4G19800, AT2G43620, AT4G19810, AT3G04720, AT4G19820, AT3G12500, AT5G24090, AT3G16920, AT3G47540, AT3G54420,
GO:0004601	AT1G05240, AT1G77490, AT3G06050, AT4G16270 AT5G22410, AT1G05250, AT2G22420, AT3G09640, AT4G21960 AT5G24070, AT1G05260, AT2G24800, AT3G11630, AT4G25980

	AT5G39580, AT1G07890, AT2G25080, AT3G17070, AT4G26010 AT5G42180, AT1G14540, AT2G31570, AT3G21770, AT4G31760 AT5G51890, AT1G30870, AT2G34060, AT3G42570, AT4G31870 AT5G64100, AT1G34510, AT2G37130, AT3G49110, AT4G32320 AT5G64110, AT1G35720, AT2G39040, AT3G49120, AT4G33420 AT5G64120, AT1G44970, AT2G41480, AT3G50990, AT4G33870, AT1G63460, AT2G43350, AT3G63080, AT4G35000, AT1G65970, AT2G43480, AT4G08390, AT4G35970, AT1G71695, AT2G48150, AT4G09010, AT5G06290, AT1G73680, AT3G01190, AT4G11290, AT5G15180, AT1G77100, AT3G03670, AT4G11600, AT5G17820,
GO:0004620	AT1G05800, AT4G10030, AT1G06800, AT4G13550, AT1G30370, AT4G16820, AT1G31480, AT4G19860, AT1G51440, AT4G37050, AT1G61850, AT4G37060, AT2G15230, AT4G37070, AT2G26560, AT5G43590, AT2G30550, , AT2G31690, AT2G39220, AT2G44810, AT3G15730, AT3G54950,
GO:0004650	AT1G02460, AT1G65570, AT2G43870, AT3G48950 AT4G32375, AT1G02790, AT1G70500, AT2G43880, AT3G57510 AT4G32380, AT1G05650, AT1G78400, AT2G43890, AT3G57790 AT4G33440, AT1G05660, AT1G80140, AT3G06770, AT3G59850 AT4G35670, AT1G10640, AT1G80170, AT3G07820, AT3G61490 AT5G14650, AT1G17150, AT2G15450, AT3G07830, AT3G62110 AT5G17200, AT1G19170, AT2G15460, AT3G07840, AT4G01890 AT5G27530, AT1G23460, AT2G15470, AT3G07850, AT4G13760 AT5G39910, AT1G43080, AT2G23900, AT3G07970, AT4G18180 AT5G41870, AT1G43090, AT2G26620, AT3G14040, AT4G20040 AT5G44830, AT1G43100, AT2G33160, AT3G15720, AT4G20050 AT5G44840, AT1G48100, AT2G40310, AT3G16850, AT4G23500 AT5G48140 AT1G56710, AT2G41850, AT3G26610, AT4G23820 AT5G49215, AT1G60590, AT2G43860, AT3G42950, AT4G32370,
GO:0004784	AT1G08830, AT1G12520, AT2G28190, AT3G10920, AT3G56350, AT4G25100, AT5G18100, AT5G23310, AT5G51100
GO:0005471	AT1G15500, AT1G80300, AT3G08580, AT4G28390, AT5G05310, AT5G13490, AT5G17400, AT5G56450
GO:0005884	AT1G03470, AT5G58320, AT1G58210, AT1G64500, AT2G04750, AT3G22790, AT3G57410, AT4G25590, AT4G26700, AT4G30160, AT5G11710, AT5G35700, AT5G48460, AT5G52360, AT5G55400,
GO:0006869	AT1G07747, AT2G14846, AT3G12545, AT3G51600 AT5G38180, AT1G19800, AT2G15050, AT3G18280, AT3G57310 AT5G38195, AT1G27950, AT2G15325, AT3G20320, AT3G59455 AT5G38197, AT1G31812,, AT2G16592, AT3G22600, AT4G08670 AT5G50800, AT1G36150, AT2G16594, AT3G29152, AT4G12825 AT5G53470, AT1G43665, AT2G18370, AT3G43720, AT4G14815 AT5G59310, AT1G43666, AT2G37870, AT3G47730, AT4G25010 AT5G59320, AT1G43667, AT2G38540, AT3G47740, AT4G27780 AT5G59330, AT1G48750, AT2G41700, AT3G47750, AT4G33355 AT5G61690, AT1G64235, AT2G44290, AT3G47760, AT5G01870 AT5G61700, AT1G65410, AT2G44300, AT3G47770, AT5G09370 AT5G61730, AT1G66850,, AT3G05420, AT3G47780, AT5G27630 AT5G61740, AT1G73780, AT3G08770, AT3G47790, AT5G38160 AT5G62065, AT2G13820, AT3G11670, AT3G51590, AT5G38170. AT5G64080
GO:0006986	AT3G10800
GO:0008234	AT1G03720, AT5G17140, AT1G37020, AT5G28235, AT1G47128, AT5G45570, AT1G55350, AT5G45890 AT1G79340, AT5G60360, AT2G07200, AT2G07240, AT2G10625, AT3G48480, AT4G08430, AT4G15880, AT5G04200, AT5G05050, AT5G17080,
GO:0008519	AT1G64780, AT2G38290, AT3G16240, AT3G24290, AT3G24300, AT4G13510, AT4G28700, AT5G47450
GO:0009326	AT5G14780

GO:0009916	AT1G32350, AT3G22360, AT3G22370, AT3G27620, AT4G05020, AT4G2226, AT5G64210
GO:0015081	AT1G08960, AT1G14660, AT1G54370, AT3G05030, AT3G06370, AT3G14070, AT4G10310, AT5G27150, AT5G55470
GO:0015200	AT3G16240, AT5G47450
GO:0015250	AT1G01620, AT2G45960, AT4G19030, AT1G17810, AT3G04090, AT4G23400, AT1G31880, AT3G06100, AT4G35100, AT1G52180, AT3G16240, AT5G18290, AT1G73190, AT3G26520, AT5G37810, AT1G80760, AT3G47440, AT5G37820, AT2G16850, AT3G53420, AT5G47450, AT2G25810, AT3G54820, AT5G60660, AT2G29870, AT3G61430, AT2G34390, AT4G00430, AT2G36830, AT4G01470, AT2G37170, AT4G10380, AT2G37180, AT4G17340, AT2G39010, AT4G18910,
GO:0015843	AT1G64780
GO:0015926	AT5G63840
GO:0016298	AT1G11090, AT1G58525, AT1G77420, AT3G55180 AT5G40990, AT1G18360, AT1G58725, AT2G26560, AT3G55190 AT5G55050, AT1G20120, AT1G59030, AT2G39220, AT3G62860 AT5G63170, AT1G20130,, AT1G59406, AT2G39400, AT4G13550, AT1G20132, AT1G71120, AT2G39410, AT4G20670, AT1G20135, AT1G71250, AT2G39420, AT4G24160, AT1G23240, AT1G73480, AT2G47630, AT5G03810, AT1G23500, AT1G73610, AT3G14225, AT5G03820, AT1G52760, AT1G75880, AT3G14820, AT5G11650, AT1G53920, AT1G75890, AT3G26430, AT5G14180, AT1G53940, AT1G75900, AT3G43550, AT5G14980, AT1G53970, AT1G75910, AT3G43570, AT5G15720, AT1G53990, AT1G75920, AT3G48090, AT5G16120, AT1G54030, AT1G75930, AT3G52430, AT5G19290,
GO:0016791	AT1G01610, AT2G34740, AT3G63320, AT4G33920 AT5G47760, AT1G02390, AT2G38110, AT3G63340, AT5G01270 AT5G57440, AT1G05000, AT2G39840, AT4G00400, AT5G02760 AT5G59220, AT1G06520, AT2G41460, AT4G01950, AT5G04120 AT5G66080, AT1G07010, AT3G02800, AT4G11040, AT5G06090, AT1G08940, AT3G05170, AT4G15030, AT5G06750, AT1G17545, AT3G11325, AT4G18140, AT5G10460, AT1G17710, AT3G11410, AT4G21670, AT5G16480, AT1G18480, AT3G11430, AT4G25840, AT5G22620, AT1G73010, AT3G16800, AT4G26080, AT5G27930, AT1G79630, AT3G17790, AT4G27800, AT5G36250, AT2G14110, AT3G23360, AT4G28400, AT5G36700, AT2G32960, AT3G50520, AT4G29530, AT5G39400, AT2G33255, AT3G58830, AT4G32950, AT5G46410,
GO:0016887	AT1G01910, AT1G10680, AT1G30410, AT1G53750 AT1G73110 AT2G13610, AT2G28290, AT2G47000, AT3G12810 AT3G24530 AT3G28600, AT3G47790, AT3G55130, AT4G02480 AT4G14700 AT4G27180, AT4G34720, AT5G03340, AT5G16930 AT5G28540 AT5G53010, AT5G61690, , AT1G01950, AT1G12430 AT1G30420 AT1G53780, AT1G73170, AT2G16510, AT2G28520 AT2G47800 AT3G13080, AT3G25620, AT3G28610, AT3G47950 AT3G56690 AT4G04180, AT4G18050, AT4G27420, AT4G35520 AT5G03910 AT5G17730, AT5G39040, AT5G53350, AT5G61700 AT1G02520 AT1G16030, AT1G30580, AT1G54350 , AT1G74310 AT2G18190 AT2G32120, AT3G01540, AT3G13090, AT3G27120 AT3G28860 AT3G48870, AT3G57300, AT4G04910, AT4G18480 AT4G27680 AT4G36280, AT5G06530, AT5G17740, AT5G40000 AT5G53540 AT5G61730, AT1G02530, AT1G17260, AT1G31770 AT1G55550 AT1G75630, AT2G18193, AT2G34560, AT3G01610 AT3G13100 AT3G27610, AT3G29800, AT3G49650, AT3G59140 AT4G05190 AT4G19210, AT4G28000, AT4G36290, AT5G06670 AT5G17750 AT5G40010, AT5G54670, AT5G61740, AT1G02890 AT1G17840 AT1G33290, AT1G56050, AT1G80350, AT2G18330 AT2G34660 AT3G03060, AT3G13640, AT3G28345, AT3G30875 AT3G50930 AT3G60160, AT4G05340, AT4G21270, AT4G28070, AT4G36580 AT5G08470, AT5G17760, AT5G42020, AT5G55520 AT5G62670, AT1G03000, AT1G18550, AT1G33360, AT1G56410 AT1G80660 AT2G18960, AT2G36910, AT3G04080, AT3G15120

	AT3G28360 AT3G42640, AT3G50940, AT3G60330, AT4G05380 AT4G24190 AT4G28620, AT4G37270, AT5G08680, AT5G18280 AT5G42270 AT5G56030, AT5G64840, AT1G04120, AT1G19100 AT1G43910 AT1G62130 , AT2G01120, AT2G21380, AT2G37010AT3G05530 AT3G17360, AT3G28380, AT3G44050, AT3G52310 AT3G60970 AT4G09140, AT4G24710, AT4G28630, AT4G37910 AT5G08690 AT5G19410, AT5G43010, AT5G57350, AT5G65930 AT1G04730 AT1G19910, AT1G45000, AT1G64110, AT2G01220 AT2G21410 AT2G37360, AT3G09440, AT3G19050, AT3G28390 AT3G45450 AT3G53230, AT3G62150, AT4G09730, AT4G24860 AT4G29040 AT4G38920, AT5G09590, AT5G19990, AT5G45930 AT5G57480 AT5G66750, AT1G05910, AT1G20060, AT1G50140 AT1G64550 AT2G01320, AT2G24520, AT2G39350, AT3G09840 AT3G19740 AT3G28415, AT3G47730, AT3G53510, AT3G62700 AT4G11730 AT4G24970, AT4G30190, AT4G39050, AT5G09930 AT5G23630 AT5G46540, AT5G57710, ATCG00120, AT1G07670 AT1G21690 AT1G51460, AT1G65070, AT2G03670, AT2G25140 AT2G39730 AT3G10350, AT3G20150, AT3G28510, AT3G47740 AT3G54540 AT3G63480, AT4G12620, AT4G25450, AT4G30250 AT4G39080 AT5G11170, AT5G23910, AT5G49840, AT5G58270 ATCG00480, AT1G07810, AT1G21730, AT1G51500, AT1G65410 AT2G07560 AT2G25170, AT2G40130, AT3G10420, AT3G21090 AT3G28520 AT3G47750, AT3G54870, AT4G00900, AT4G14150 AT4G25750 AT4G30490, AT4G39850, AT5G11200, AT5G24020 AT5G50780 AT5G58290, ATCG01280, AT1G09080, AT1G27940 AT1G51650 AT1G67120, AT2G07680, AT2G25530, AT2G41700 AT3G10670 AT3G21250, AT3G28540, AT3G47760, AT3G55090 AT4G01820 AT4G14330, AT4G25835, AT4G32208, AT5G02370 AT5G13130 AT5G27000, AT5G51070, AT5G60730, ATMG01190 AT1G09100 AT1G28010, AT1G53270, AT1G67940, AT2G07698 AT2G25610 AT2G45500, AT3G12020, AT3G23255, AT3G28570 AT3G47770 AT3G55100, AT4G01830, AT4G14670, AT4G25960 AT4G32530, AT5G02490, AT5G13580, AT5G27550, AT5G52860 AT5G60740, AT1G10130, AT1G30400, AT1G53390, AT1G70610 AT2G11890 AT2G27600, AT2G46620, AT3G12580, AT3G23670 AT3G28580 AT3G47780, AT3G55110, AT4G02460, AT4G14675 AT4G26660 AT4G33460, AT5G02500, AT5G15450, AT5G27950 AT5G52882 AT5G60790, ,
GO:0019137	AT1G47600, AT1G51470, AT2G44490, AT3G09260, AT5G25980, AT5G26000, AT5G48375
GO:0031176	AT1G10050, AT1G58370, AT2G14690, AT4G00890, AT4G08160, AT4G33810, AT4G33820
GO:0034620	AT1G09080, AT5G28540, AT1G16030, AT5G42020, AT1G56410, AT5G62390, AT2G26150, AT2G32120, AT3G09440, AT3G12580, AT4G02690, AT4G14730, AT4G32208, AT4G37910, AT5G02490, AT5G02500, AT5G09590,
GO:0042026	AT1G09080, AT3G13470, AT5G42020, AT1G16030, AT3G13860, AT5G48030, AT1G26230, AT3G17830, AT5G56500, AT1G28210, AT3G23990, AT5G59610, AT1G55490, AT4G32208, AT1G56410, AT4G37910, AT1G80030, AT4G39960, AT2G22360, AT5G02490, AT2G25140, AT5G02500, AT2G28000, AT5G09590, AT2G32120, AT5G15450, AT2G33210, AT5G18140, AT3G09440, AT5G18820, AT3G12580, AT5G28540,
GO:0042803	AT1G01360, AT1G13980, AT1G45474, AT1G74660 AT2G19570 AT2G35940, AT2G47220, AT3G13290, AT3G49250 AT3G63480 AT4G18880, AT4G27920, AT5G06200, AT5G18830 AT5G45870 AT5G61850, AT1G01510, AT1G14440, AT1G49950 AT1G75240 AT2G21240, AT2G36100, AT2G47430, AT3G13300 AT3G49850 AT4G00570, AT4G19003, AT4G30260, AT5G06839 AT5G24020 AT5G46790, AT5G62380, AT1G01760, AT1G14687 AT1G51800 AT1G77120, AT2G21280, AT2G37630, AT3G01470 AT3G13672 AT3G50870, AT4G00650, AT4G23690, AT4G30790 AT5G10150 AT5G24800, AT5G47850, AT5G63570, AT1G02280 AT1G15750 AT1G55250, AT1G80560 , AT2G24630 AT2G37678, AT3G01670 AT3G15000 AT3G50890, AT4G01026, AT4G24540, AT4G31590 AT5G10870 AT5G27380, AT5G48840, AT5G63860 AT1G03475, AT1G17455, AT1G62360, AT2G02540, AT2G26040, AT2G38310, AT3G01680, AT3G21630, AT3G51960, AT4G02500, AT4G24660, AT4G34710, AT5G11950, AT5G35840, AT5G49020, AT5G65310 AT1G04400, AT1G17840, AT1G67930, AT2G03220, AT2G27230, AT2G38700, AT3G03890, AT3G26400, AT3G54220, AT4G07960, AT4G24840, AT4G35540, AT5G11980, AT5G37740, AT5G50210, AT5G65410

	AT1G05577, AT1G18835, AT1G68880, AT2G04030, AT2G27370, AT2G39180, AT3G06483, AT3G28180, AT3G55950, AT4G08920, AT4G25530, AT4G35630, AT5G12330, AT5G38530, AT5G50960, AT5G65930 AT1G08060, AT1G19100, AT1G69370, AT2G06255, AT2G28150, AT2G40080, AT3G06930, AT3G28917, AT3G57230, AT4G14180, AT4G26455, AT4G38710, AT5G13220, AT5G39610, AT5G53160, AT5G67580 AT1G08320, AT1G26110, AT1G69600, AT2G13560, AT2G28900, AT2G40330, AT3G06960, AT3G28920, AT3G58840, AT4G16160, AT4G26670, AT4G38900, AT5G13570, AT5G39760, AT5G54770, ATOEP16-2 AT1G08370, AT1G26540, AT1G71930, AT2G15290, AT2G29950, AT2G42210, AT3G07330, AT3G29200, AT3G59420, AT4G16250, AT4G26740, AT4G39400, AT5G13790, AT5G40610, AT5G55200, ATOEP16-4 AT1G09570, AT1G30970, AT1G72630, AT2G16500, AT2G33430, AT2G42870, AT3G07780, AT3G30530, AT3G61600, AT4G16780, AT4G26750, AT5G01410, AT5G15210, AT5G42520, AT5G56210, AT1G11430, AT1G32580, AT1G73000, AT2G18350, AT2G33550, AT2G44950, AT3G09260, AT3G46110, AT3G61950, AT4G17870, AT4G26780, AT5G02200, AT5G15290, AT5G42780, AT5G57590, AT1G13020, AT1G35720, AT1G73430, AT2G18790, AT2G33580, AT2G46260, AT3G09780, AT3G48090, AT3G62300, AT4G18130, AT4G26820, AT5G03150, AT5G17710, AT5G45330, AT5G59790, AT1G13600, AT1G36390, AT1G74230, AT2G18840, AT2G35240, AT2G46810, AT3G11550, AT3G48730, AT3G62880, AT4G18620, AT4G27420, AT5G03680, AT5G17800, AT5G45860, AT5G60480,
GO:0045551	AT1G30760, AT1G72680, AT2G21730, AT2G21890, AT2G34790, AT3G19450, AT4G27250, AT4G34230, AT4G37970, AT4G37980, AT4G37990, AT4G39330
GO:0052861	AT1G18310, AT1G64760, AT5G15870
GO:0055044	AT3G15380
GO:0072488	AT1G64780, AT2G38290, AT3G24290, AT3G24300, AT4G13510, AT4G28700
GO:0090353	AT5G06860, AT5G06870

Table 5. The lenient number of genes matching the annotated GO codes from *A. thaliana*, for each taxon is reported.

Go Code Definition	**GO Code Label**	**GO Code**	*Cephalotus follicularis*	*Aldrovanda vesiculosa*	*Dionaea muscipula*	*Drosera spatulata*	*Utricularia reniformis*	*Utricularia gibba*	*Genlisea aurea*	*Beta vulgaris*	*Populus trichocarpa*	*Dorcoceras*	*Erythranthe guttata*	*Glycine soja*	*Olea europaea*	*Phtheirospermum*	*Oryza sativa*	*Coffea canephora*	*Sesamum indicum*	*Spinacia oleracea*	*Striga asiatica*	*Solanum lycopersicum*	*Salvia splendens*	*Handroanthus*	*Chenopodium quinoa*
cyclic-nucleotide phosphodiesterase activity	1	GO:0004112	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Serine-type carboxypeptidase activity	2	GO:0004185	38	19	27	31	27	21	29	40	66	34	46	149	93	53	21	40	58	34	46	65	66	64	101
Aspartic-type endopeptidase activity	3	GO:0004190	32	46	23	22	27	16	32	33	124	19	38	131	68	21	15	33	52	17	24	58	48	43	52
Fructose-bisphosphate aldolase activity	4	GO:0004332	42	116	63	62	39	93	52	72	90	33	54	147	90	69	72	63	75	54	78	71	150	80	104
Glutathione transferase activity	5	GO:0004364	79	18	38	24	13	11	16	38	99	31	20	166	39	20	6	65	55	46	25	87	30	24	148
Endonuclease activity	6	GO:0004519	25	7	5	6	14	12	13	9	26	8	12	40	20	18	2	11	24	12	13	22	13	18	24
Ribonuclease activity	7	GO:0004540	31	10	8	10	11	6	10	12	30	7	21	93	26	24	9	11	21	13	15	30	22	16	35
Alpha-galactosidase	8	GO:0004557	5	6	9	6	14	20	10	6	12	3	4	17	17	13	6	6	15	8	3	8	15	9	13

activity																								
Beta-galactosidase activity	9	GO:0004565	67	121	88	57	98	56	105	69	147	50	147	253	122	125	26	119	122	75	63	127	240	98	132
Chitinase activity	10	GO:0004568	16	16	8	6	14	5	5	7	33	11	22	36	32	7	21	10	15	8	9	20	32	19	15
Peroxidase activity	11	GO:0004601	106	100	63	68	100	86	69	146	178	70	122	238	259	129	95	131	157	101	105	180	197	129	203
Phospholipase activity	12	GO:0004620	23	8	8	6	8	2	6	20	76	11	23	31	39	20	7	26	25	14	16	24	25	33	25
Polygalacturonase activity	13	GO:0004650	46	24	18	18	24	34	23	22	139	27	32	215	58	58	7	36	53	29	27	64	53	18	51
Superoxide dismutase activity	14	GO:0004784	27	16	24	23	16	24	19	33	95	22	41	97	72	23	19	27	37	51	26	52	36	46	99
ATP:ADP antiporter activity	15	GO:0005471	47	45	28	18	17	26	16	21	128	28	43	107	72	46	17	35	38	12	41	50	91	41	45
Actin filament	16	GO:0005884	80	127	76	59	94	73	64	107	185	63	109	584	291	107	41	62	107	106	92	195	202	53	193
Lipid transport	17	GO:0006869	44	30	19	16	20	20	22	40	86	26	43	150	64	27	9	31	87	44	41	40	42	36	96
response to unfolded protein	18	GO:0006986	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cysteine-type peptidase activity	19	GO:0008234	12	23	11	10	22	11	11	11	50	10	26	25	35	17	9	13	22	11	18	28	38	16	22
Ammonium transmembrane	20	GO:0008519	20	31	15	21	25	34	11	31	52	18	39	45	45	23	32	28	35	26	17	27	64	34	60

transport																								
Formate dehydrogenase complex	21	GO:0009326	6	3	6	6	3	3	3	3	6	3	3	18	3	3	6	6	6	3	3	3	9	3	6
Alternative oxidase activity	22	GO:0009916	20	20	21	21	12	18	9	14	30	11	21	28	51	19	23	20	31	23	26	25	47	28	41
Sodium ion transmembrane transporter activity	23	GO:0015081	51	19	9	12	3	20	26	13	122	23	30	63	189	28	27	37	47	25	38	76	47	25	28
Aquaporin, Methylammonium transmembrane transporter activity	24	GO:0015200	4	0	0	0	2	2	2	4	8	2	7	12	11	4	6	4	6	4	2	8	18	8	8
water channel activity	25	GO:0015250	230	308	177	170	249	163	126	187	389	240	230	632	530	335	214	177	299	222	288	313	679	358	307
methylammonium transport	26	GO:0015843	3	5	3	3	3	6	1	5	6	2	6	5	5	3	3	4	5	4	3	3	8	4	9
Glucosidase activity	27	GO:0015926	2	4	2	2	2	4	2	2	4	2	6	2	4	2	2	4	2	2	2	2	4	0	4
lipase activity	28	GO:0016298	49	18	15	10	23	22	10	33	81	22	34	76	62	28	7	39	36	24	23	32	33	41	49
phosphatase activity	29	GO:0016791	69	74	58	48	59	53	56	77	215	57	99	252	295	69	54	70	121	94	59	120	153	81	192
ATPase	30	GO:0016887	14	15	862	882	970	903	791	12	30	10	14	31	23	14	941	991	17	12	11	15	25	12	25

activity			83	71						76	03	48	19	98	94	00			67	41	54	14	74	09	64
thioglucosidase activity	31	GO:0019137	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Endo-1,4-beta-xylanase activity	32	GO:0031176	4	4	4	4	2	0	0	16	10	2	2	32	8	2	0	4	2	8	0	10	0	7	20
heat shock protein activity	33	GO:0034620	79	160	101	118	119	154	104	139	210	145	165	396	207	156	161	138	138	157	218	161	72	132	345
heat shock protein activity	34	GO:0042026	132	235	176	161	159	207	139	221	347	253	218	539	310	223	206	202	195	226	267	268	180	203	523
protein homodimerization activity	35	GO:0042803	317	291	177	174	238	248	195	299	682	218	340	983	694	315	222	252	568	293	246	445	598	354	520
cinnamyl-alcohol dehydrogenase activity	36	GO:0045551	38	8	10	9	27	16	14	17	77	18	24	60	59	31	7	53	37	21	20	30	97	35	45
Glucan endo-1,3-beta-glucanase activity	37	GO:0052861	5	2	3	4	0	2	3	6	6	3	3	8	3	3	4	3	3	7	4	5	10	3	8
symplast	38	GO:0055044	1	2	1	1	1	1	1	1	1	1	1	4	3	1	2	2	1	1	2	1	3	0	4
Ammonium transmembrane transport	39	GO:0072488	16	31	15	21	23	2	9	27	44	16	32	33	34	19	26	24	29	22	15	19	46	26	52
polygalacturonase inhibit	40	GO:0090353	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

or activity																							

Table 6. The conservative number of genes matching the annotated GO codes from *A. thaliana*, for each taxon is reported.

Go Code Definition	**GO Code Label**	**GO Code**	*Cephalotus follicularis*	*Aldrovanda vesiculosa*	*Dionaea muscipula*	*Drosera spatulata*	*Utricularia reniformis*	*Utricularia gibba*	*Genlisea aurea*	*Beta vulgaris*	*Populus trichocarpa*	*Dorcoceras hygrometricum*	*Erythranthe guttata*	*Glycine soja*	*Olea europaea*	*Phtheirospermum japonicum*	*Oryza sativa*	*Coffea canephora*	*Sesamum indicum*	*Spinacia oleracea*	*Striga asiatica*	*Solanum lycopersicum*	*Salvia splendens*	*Handroanthus impetiginosus*	*Chenopodium quinoa*
cyclic-nucleotide phosphodiesterase activity	1	GO:0004112	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Serine-type carboxypeptidase activity	2	GO:0004185	7	0	8	5	5	4	0	7	23	3	8	17	8	7	2	3	9	4	2	9	6	7	15
Aspartic-type endopeptidase activity	3	GO:0004190	5	4	0	2	0	1	0	8	66	11	15	44	9	5	0	14	22	2	13	16	6	25	4
Fructose-bisphosphate aldolase activity	4	GO:0004332	34	80	53	53	48	74	42	40	64	23	42	118	69	54	54	51	58	38	59	57	111	53	78
Glutathione	5	GO	8	1	0	0	1	4	2	4	13	3	4	35	5	3	0	4	10	2	4	11	3	4	12

transferase activity		:0004364																							
Endonuclease activity	6	GO:0004519	4	1	2	3	1	1	0	2	2	1	2	18	4	2	0	3	2	2	2	3	2	6	6
Ribonuclease activity	7	GO:0004540	11	9	7	6	12	5	4	10	11	6	11	37	14	6	3	5	9	6	13	10	12	7	26
Alpha-galactosidase activity	8	GO:0004557	2	2	2	1	2	0	1	0	5	0	2	0	4	3	0	2	1	4	0	0	1	4	4
Beta-galactosidase activity	9	GO:0004565	8	7	2	2	3	0	12	1	13	3	5	24	13	10	0	3	8	1	1	5	14	8	2
Chitinase activity	10	GO:0004568	5	2	2	2	8	0	0	0	4	2	0	10	12	0	0	0	1	0	2	1	4	7	0
Peroxidase activity	11	GO:0	34	44	29	32	47	12	19	39	59	17	56	67	57	32	40	27	58	25	36	49	56	34	66

		0046 01																							
Phospholipase activity	12	GO:0004620	3	1	1	1	2	2	1	1	5	2	4	9	11	4	2	3	3	2	2	7	3	3	4
Polygalacturonase activity	13	GO:0004650	5	0	0	0	0	2	1	0	4	3	3	5	4	3	0	1	1	0	0	0	2	0	0
Superoxide dismutase activity	14	GO:0004784	2	0	1	0	0	0	4	3	7	0	3	2	9	1	1	4	5	2	0	6	2	9	8
ATP:ADP antiporter activity	15	GO:0005471	19	33	20	6	7	11	9	10	97	22	20	68	17	16	10	16	25	7	27	34	54	30	20
Actin filament	16	GO:0005884	13	2	1	2	3	8	11	5	44	8	12	55	53	11	0	5	12	5	4	46	3	13	22
Lipid transport	17	GO:000 0	11	0	8	1	1	1	10	13	8	7	1	3	10	6	1	7	68	0	7	15	9	16	44

		6869																							
Response to unfolded protein	18	GO:0006986	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cysteine-type peptidase activity	19	GO:0008234	9	6	8	3	12	0	0	9	45	5	24	12	27	9	0	9	18	9	9	18	23	9	18
Ammonium transmembrane transport	20	GO:0008519	13	15	5	4	13	15	6	5	28	10	22	25	31	6	6	11	16	4	12	13	39	12	13
Formate dehydrogenase complex	21	GO:0009326	6	3	6	6	3	3	3	3	3	3	3	18	3	3	6	6	6	3	3	3	9	3	6
Alternative oxidase activity	22	GO:0009916	4	4	6	6	2	0	2	0	8	1	6	4	5	1	2	1	6	2	4	7	6	4	5
Sodium ion transmembrane transport	23	GO:0015	20	5	2	2	2	5	10	6	43	4	16	30	64	11	7	15	18	6	6	28	5	6	12

er activity		0081																							
Aquaporin, Methylammonium transmembrane transporter activity	24	GO:0015200	4	0	0	0	0	0	1	2	7	1	3	8	10	3	2	2	3	2	2	6	9	2	4
water channel activity	25	GO:0015250	104	74	61	40	88	56	33	77	198	104	127	313	240	111	78	61	170	87	144	153	337	156	126
methylammonium transport	26	GO:0015843	2	4	2	1	4	5	1	3	4	1	5	5	4	1	2	2	4	2	2	2	6	2	4
Glucosidase activity	27	GO:0015926	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
lipase activity	28	GO:0016298	3	5	6	2	6	5	3	4	10	3	4	9	13	10	2	4	6	2	6	3	4	8	5
phosphatase activity	29	GO:0016	20	14	15	7	12	11	11	20	65	8	19	53	45	14	9	18	26	18	16	17	26	19	54

		7911																							
ATPase activity	30	GO:0016887	440	589	341	345	665	448	383	550	1249	503	681	1408	1069	531	423	451	902	469	487	652	1169	614	1182
Thiogluc osidase activity	31	GO:0019137	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Endo-1,4-beta-xylanase activity	32	GO:0031176	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
heat shock protein activity	33	GO:0034620	58	101	69	85	130	79	68	70	147	64	102	179	143	109	78	75	86	67	70	116	10	90	154
heat shock protein activity	34	GO:0042026	87	145	109	103	154	104	80	102	220	92	120	236	189	138	102	102	104	85	102	151	69	110	199
protein homodi merizati on activity	35	GO:00428	90	68	39	25	75	34	48	66	202	48	74	254	135	45	47	77	109	77	50	112	133	80	141

		03																							
cinnamyl-alcohol dehydrogenase activity	36	GO:0045551	8	3	3	3	1	0	1	4	23	2	3	9	11	2	0	4	5	4	3	4	2	4	9
Glucan endo-1,3-beta-glucanase activity	37	GO:0052861	2	0	0	0	0	0	0	0	2	0	2	0	0	0	0	0	0	0	0	2	2	0	0
symplast	38	GO:0055044	1	0	0	0	0	0	0	1	1	0	0	4	0	0	0	1	0	0	0	1	0	0	0
Ammonium transmembrane transport	39	GO:0072488	9	15	5	4	13	15	5	3	21	9	19	17	21	3	4	9	13	2	10	7	30	10	9
Polygalacturonase inhibitor activity	40	GO:0090353	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 7. The lenient number of gene matches using a comparative phylogenic analysis excluding closely related taxa of *A. thaliana*.

Go Code Definition	GO Code Label	GO Code	*Aldrovanda vesiculosa*	*Dionaea muscipula*	*Drosera spatulata*	*Utricularia reniformis*	*Utricularia gibba*	*Genlisea aurea*	*Beta vulgaris*	*Dorcoceras hygrometricum*	*Erythranthe guttata*	*Olea europaea*	*Phtheirospermum japonicum*	*Oryza sativa*	*Coffea canephora*	*Sesamum indicum*	*Spinacia oleracea*	*Striga asiatica*	*Solanum lycopersicum*	*Salvia splendens*	*Handroanthus impetiginosus*	*Chenopodium quinoa*
cyclic-nucleotide phosphodiesterase activity	1	GO:0004112	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Serine-type carboxypeptidase activity	2	GO:0004185	19	27	31	27	21	29	40	34	46	93	53	21	40	58	34	46	65	66	64	101
Aspartic-type endopeptidase activity	3	GO:0004190	46	23	22	27	16	32	33	19	38	68	21	15	33	52	17	24	58	48	43	52
Fructose-bisphosphate aldolase activity	4	GO:0004332	116	63	62	39	93	52	72	33	54	90	69	72	63	75	54	78	71	150	80	104
Glutathione transferase activity	5	GO:0004364	18	38	24	13	11	16	38	31	20	39	20	6	65	55	46	25	87	30	24	148
Endonuclease activity	6	GO:0004519	7	5	6	14	12	13	9	8	12	20	18	2	11	24	12	13	22	13	18	24
Ribonuclease activity	7	GO:0004540	10	8	10	11	6	10	12	7	21	26	24	9	11	21	13	15	30	22	16	35
Alpha-galactosidase activity	8	GO:0004557	6	9	6	14	20	10	6	3	4	17	13	6	6	15	8	3	8	15	9	13
Beta-galactosidase activity	9	GO:0004565	121	88	57	98	56	105	69	50	147	122	125	26	119	122	75	63	127	240	98	132
Chitinase activity	10	GO:0004568	16	8	6	14	5	5	7	11	22	32	7	21	10	15	8	9	20	32	19	15
Peroxidase activity	11	GO:0004601	100	63	68	100	86	69	146	70	122	259	129	95	131	157	101	105	180	197	129	203
Phospholipase activity	12	GO:0004620	8	8	6	8	2	6	20	11	23	39	20	7	26	25	14	16	24	25	33	25

Polygalacturonase activity	13	GO:0004650	24	18	18	24	34	23	22	27	32	58	58	7	36	53	29	27	64	53	18	51
Superoxide dismutase activity	14	GO:0004784	16	24	23	16	24	19	33	22	41	72	23	19	27	37	51	26	52	36	46	99
ATP:ADP antiporter activity	15	GO:0005471	45	28	18	17	26	16	21	28	43	72	46	17	35	38	12	41	50	91	41	45
Actin filament	16	GO:0005884	127	76	59	94	73	64	107	63	109	291	107	41	62	107	106	92	195	202	53	193
Lipid transport	17	GO:0006869	30	19	16	20	20	22	40	26	43	64	27	9	31	87	44	41	40	42	36	96
Response to unfolded protein	18	GO:0006986	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cysteine-type peptidase activity	19	GO:0008234	23	11	10	22	11	11	11	10	26	35	17	9	13	22	11	18	28	38	16	22
Ammonium transmembrane transport	20	GO:0008519	31	15	21	25	34	11	31	18	39	45	23	32	28	35	26	17	27	64	34	60
Formate dehydrogenase complex	21	GO:0009326	3	6	6	3	3	3	3	3	3	3	3	6	6	6	3	3	3	9	3	6
Alternative oxidase activity	22	GO:0009916	20	21	21	12	18	9	14	11	21	51	19	23	20	31	23	26	25	47	28	41
Sodium ion transmembrane transporter activity	23	GO:0015081	19	9	12	3	20	26	13	23	30	189	28	27	37	47	25	38	76	47	25	28
Aquaporin, Methylammonium transmembrane transporter activity	24	GO:0015200	0	0	0	2	2	2	4	2	7	11	4	6	4	6	4	2	8	18	8	8
water channel activity	25	GO:0015250	308	177	170	249	163	126	187	240	230	530	335	214	177	299	222	288	313	679	358	307
methylammonium transport	26	GO:0015843	5	3	3	3	6	1	5	2	6	5	3	3	4	5	4	3	3	8	4	9
Glucosidase activity	27	GO:0015926	4	2	2	2	4	2	2	2	6	4	2	2	4	2	2	2	2	4	0	4
lipase activity	28	GO:0016298	18	15	10	23	22	10	33	22	34	62	28	7	39	36	24	23	32	33	41	49
phosphatase activity	29	GO:0016791	74	58	48	59	53	56	77	57	99	295	69	54	70	121	94	59	120	153	81	192
ATPase activity	30	GO:0016887	1571	862	882	970	903	791	1276	1048	1419	2394	1400	941	991	1767	1241	1154	1514	2574	1209	2564
Thioglucosidase activity	31	GO:0019137	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Endo-1,4-beta-	32	GO:0031176	4	4	4	2	0	0	16	2	2	8	2	0	4	2	8	0	10	0	7	20

xylanase activity																					
heat shock protein activity	33	GO:0034620	160	101	118	119	154	104	139	145	165	207	156	161	138	138	157	218	161	72	132	345
heat shock protein activity	34	GO:0042026	235	176	161	159	207	139	221	253	218	310	223	206	202	195	226	267	268	180	203	523
protein homodimerization activity	35	GO:0042803	291	177	174	238	248	195	299	218	340	694	315	222	252	568	293	246	445	598	354	520
cinnamyl-alcohol dehydrogenase activity	36	GO:0045551	8	10	9	27	16	14	17	18	24	59	31	7	53	37	21	20	30	97	35	45
Glucan endo-1,3-beta-glucanase activity	37	GO:0052861	2	3	4	0	2	3	6	3	3	3	3	4	3	3	7	4	5	10	3	8
symplast	38	GO:0055044	2	1	1	1	1	1	1	1	1	3	1	2	2	1	1	2	1	3	0	4
Ammonium transmembrane transport	39	GO:0072488	31	15	21	23	2	9	27	16	32	34	19	26	24	29	22	15	19	46	26	52
Polygalacturonase inhibitor activity	40	GO:0090353	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8. The conservative number of gene matches using a comparative phylogenic analysis excluding closely related taxa of *A. thaliana*.

Go Code Definition	**GO Code Label**	**GO Code**	*Aldrovanda vesiculosa*	*Dionaea muscipula*	*Drosera spatulata*	*Utricularia reniformis*	*Utricularia gibba*	*Genlisea aurea*	*Beta vulgaris*	*Dorcoceras hygrometricum*	*Erythranthe guttata*	*Olea europaea*	*Phtheirospermum japonicum*	*Oryza sativa*	*Coffea canephora*	*Sesamum indicum*	*Spinacia oleracea*	*Striga asiatica*	*Solanum lycopersicum*	*Salvia splendens*	*Handroanthus impetiginosus*	*Chenopodium quinoa*
cyclic-nucleotide phosphodiesterase activity	1	GO:0004112	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Serine-type carboxypeptidase activity	2	GO:0004185	0	8	5	5	4	0	7	3	8	8	7	2	3	9	4	2	9	6	7	15
Aspartic-type endopeptidase activity	3	GO:0004190	4	0	2	0	1	0	8	11	15	9	5	0	14	22	2	13	16	6	25	4
Fructose-bisphosphate aldolase activity	4	GO:0004332	80	53	53	48	74	42	40	23	42	69	54	54	51	58	38	59	57	111	53	78
Glutathione transferase activity	5	GO:0004364	1	0	0	1	4	2	4	3	4	5	3	0	4	10	2	4	11	3	4	12
Endonuclease activity	6	GO:0004519	1	2	3	1	1	0	2	1	2	4	2	0	3	2	2	2	3	2	6	6
Ribonuclease activity	7	GO:0004540	9	7	6	12	5	4	10	6	11	14	6	3	5	9	6	13	10	12	7	26
Alpha-galactosidase activity	8	GO:0004557	2	2	1	2	0	1	0	0	2	4	3	0	2	1	4	0	0	1	4	4
Beta-galactosidase activity	9	GO:0004565	7	2	2	3	0	12	1	3	5	13	10	0	3	8	1	1	5	14	8	2
Chitinase activity	10	GO:0004568	2	2	2	8	0	0	0	2	0	12	0	0	0	1	0	2	1	4	7	0
Peroxidase activity	11	GO:0004601	44	29	32	47	12	19	39	17	56	57	32	40	27	58	25	36	49	56	34	66
Phospholipase activity	12	GO:0004620	1	1	1	2	2	1	1	2	4	11	4	2	3	3	2	2	7	3	3	4
Polygalacturonase activity	13	GO:0004650	0	0	0	0	2	1	0	3	3	4	3	0	1	1	0	0	0	2	0	0
Superoxide	14	GO:0004784	0	1	0	0	0	4	3	0	3	9	1	1	4	5	2	0	6	2	9	8

dismutase activity																						
ATP:ADP antiporter activity	15	GO:0005471	33	20	6	7	11	9	10	22	20	17	16	10	16	25	7	27	34	54	30	20
Actin filament	16	GO:0005884	2	1	2	3	8	11	5	8	12	53	11	0	5	12	5	4	46	3	13	22
Lipid transport	17	GO:0006869	0	8	1	1	1	10	13	7	1	10	6	1	7	68	0	7	15	9	16	44
Response to unfolded protein	18	GO:0006986	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cysteine-type peptidase activity	19	GO:0008234	6	8	3	12	0	0	9	5	24	27	9	0	9	18	9	9	18	23	9	18
Ammonium transmembrane transport	20	GO:0008519	15	5	4	13	15	6	5	10	22	31	6	6	11	16	4	12	13	39	12	13
Formate dehydrogenase complex	21	GO:0009326	3	6	6	3	3	3	3	3	3	3	3	6	6	6	3	3	3	9	3	6
Alternative oxidase activity	22	GO:0009916	4	6	6	2	0	2	0	1	6	5	1	2	1	6	2	4	7	6	4	5
Sodium ion transmembrane transporter activity	23	GO:0015081	5	2	2	2	5	10	6	4	16	64	11	7	15	18	6	6	28	5	6	12
Aquaporin, Methylammonium transmembrane transporter activity	24	GO:0015200	0	0	0	0	0	1	2	1	3	10	3	2	2	3	2	2	6	9	2	4
water channel activity	25	GO:0015250	74	61	40	88	56	33	77	104	127	240	111	78	61	170	87	144	153	337	156	126
methylammonium transport	26	GO:0015843	4	2	1	4	5	1	3	1	5	4	1	2	2	4	2	2	2	6	2	4
Glucosidase activity	27	GO:0015926	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
lipase activity	28	GO:0016298	5	6	2	6	5	3	4	3	4	13	10	2	4	6	2	6	3	4	8	5
phosphatase activity	29	GO:0016791	14	15	7	12	11	11	20	8	19	45	14	9	18	26	18	16	17	26	19	54
ATPase activity	30	GO:0016887	589	341	345	665	448	383	550	503	681	1069	531	423	451	902	469	487	652	1169	614	1182
Thioglucosidase activity	31	GO:0019137	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Endo-1,4-beta-xylanase activity	32	GO:0031176	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

heat shock protein activity	33	GO:0034620	101	69	85	130	79	68	70	64	102	143	109	78	75	86	67	70	116	10	90	154
heat shock protein activity	34	GO:0042026	145	109	103	154	104	80	102	92	120	189	138	102	102	104	85	102	151	69	110	199
protein homodimerization activity	35	GO:0042803	68	39	25	75	34	48	66	48	74	135	45	47	77	109	77	50	112	133	80	141
cinnamyl-alcohol dehydrogenase activity	36	GO:0045551	3	3	3	1	0	1	4	2	3	11	2	0	4	5	4	3	4	2	4	9
Glucan endo-1,3-beta-glucanase activity	37	GO:0052861	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	2	0	0
symplast	38	GO:0055044	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0
Ammonium transmembrane transport	39	GO:0072488	15	5	4	13	15	5	3	9	19	21	3	4	9	13	2	10	7	30	10	9
Polygalacturonase inhibitor activity	40	GO:0090353	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

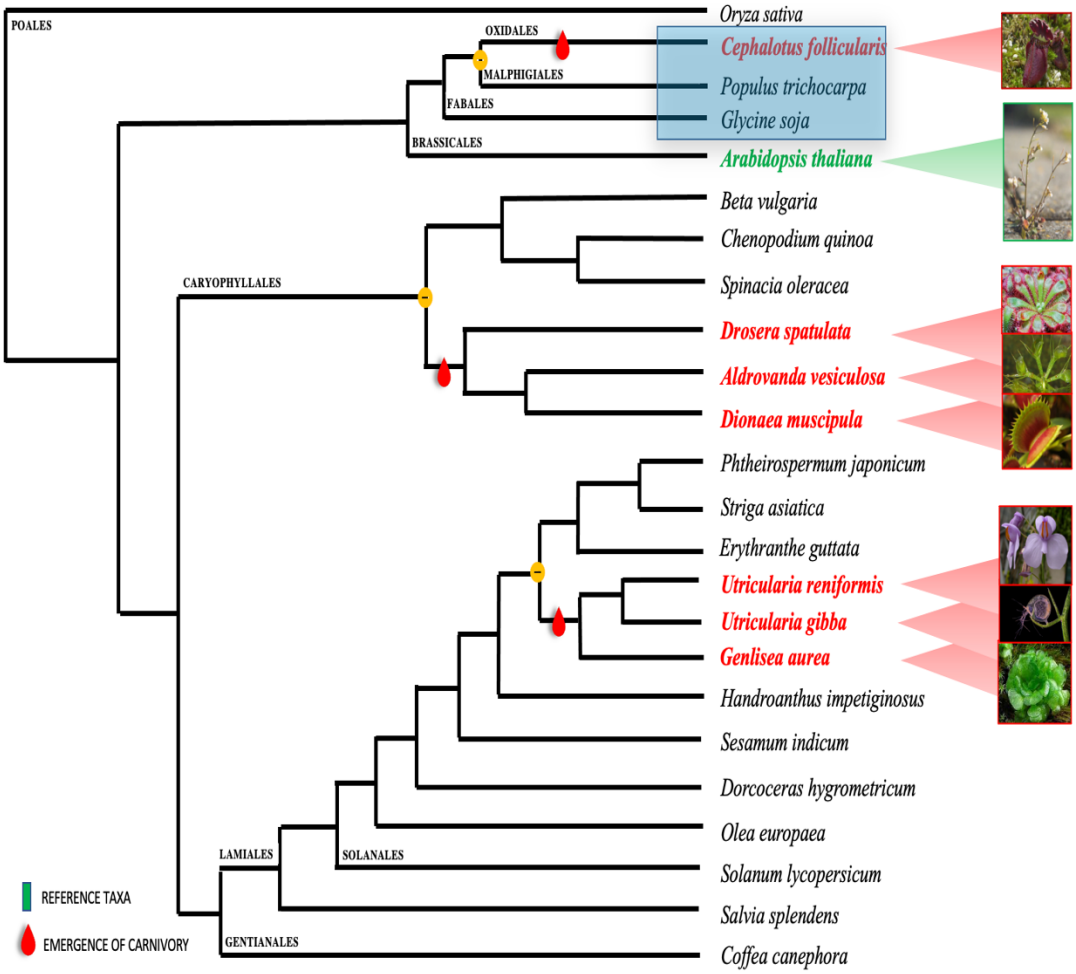

Figure 1. A neighbor-joining cladogram depicting the samples used in this study. Among these samples, carnivory evolved independently at least three separate times. The emergence of carnivorous plant species is represented by the red droplet. To the right are picture representations, indicating the various types of distinct trapping systems. The green text indicates the reference taxon, *A. thaliana*, against which all taxa were compared. A yellow circle was drawn around the advent of carnivory in plants were genes for carnivory were projected to contract or expand. The taxa highlighted in blue were removed to test for phylogenic effects.

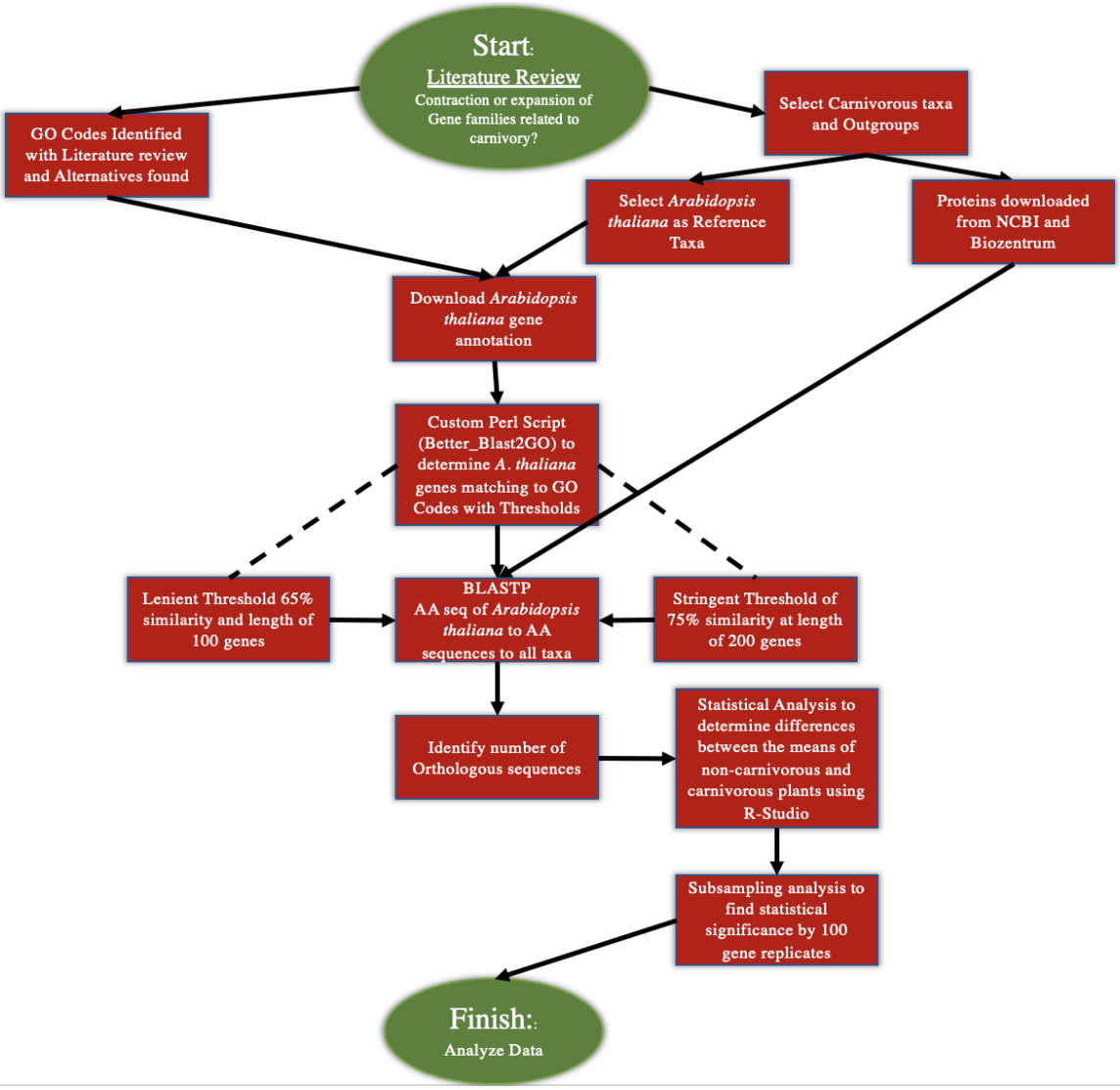

Figure 2. Flow chart of methods for genomic analysis.
This flowchart outlines the steps that were taken to determine what genes were associated with carnivorous species.

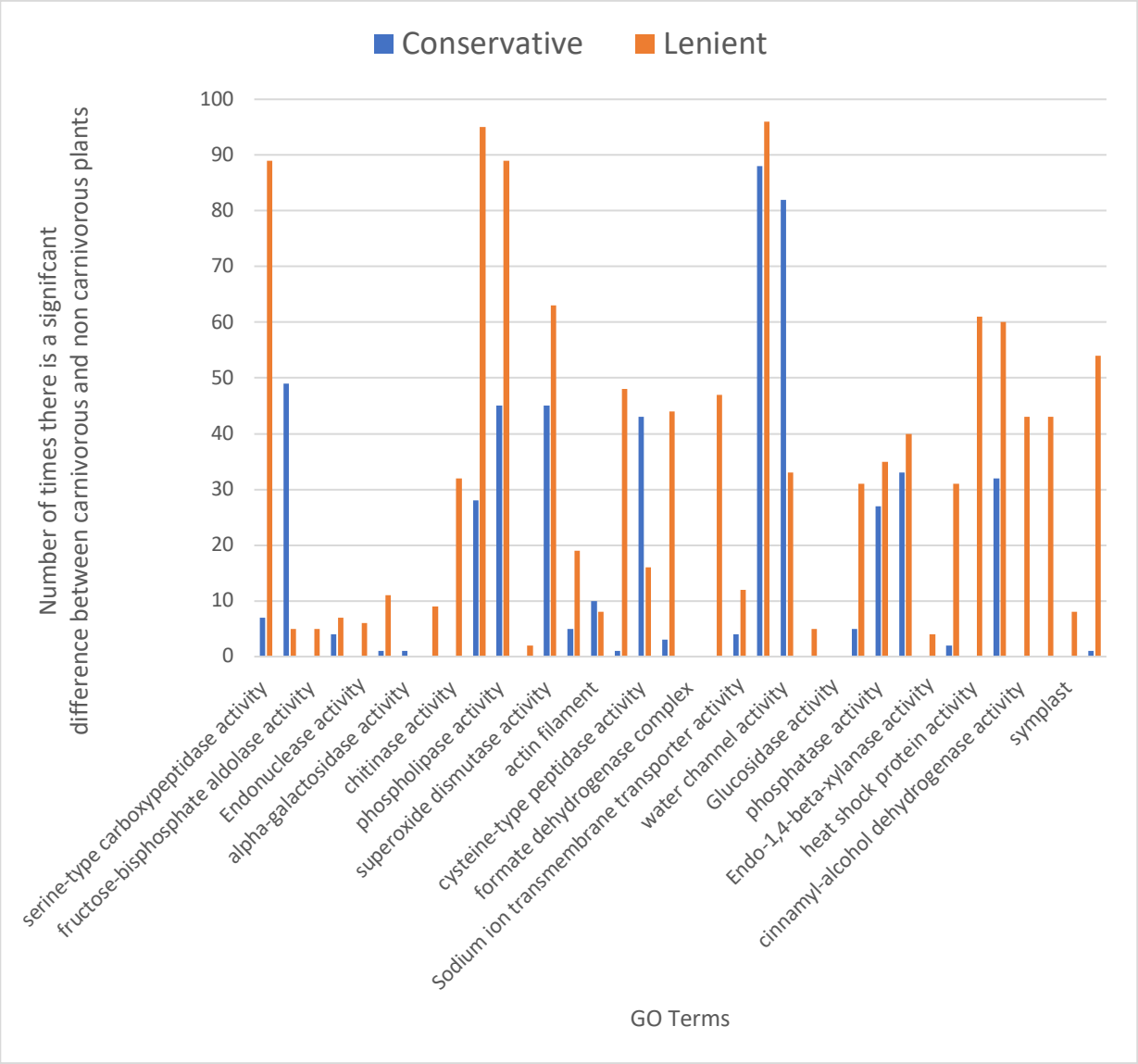

Figure 3. The subsampling method results for each GO term and the number of times significant differences found between carnivorous and non-carnivorous taxa.

A subsampling strategy was utilized to examine statistical significance across the difference in the number of gene hits identified in the BLASTP test between carnivorous and non-carnivorous plants. Using 100 gene replicates with replacements, the number of gene matches across 40 gene-related activities was examined. The lenient (green) and conservative (red) criterion for percent similarity and gene lengths are shown across each of the following GO terms. The analysis's selected GO codes define the function of genes on the x-axis. The y-axis displays the number of times the difference between carnivorous and non-carnivorous taxa was statistically significant based on p-value >0.05.