

Conventional and Novel G γ Protein Families Constitute the Heterotrimeric G-Protein Signaling Network in Soybean

Swarup Roy Choudhury, Naveen C. Bisht^{‡a}, Rheannon Thompson^{‡b}, Oleg Todorov, Sona Pandey*

Donald Danforth Plant Science Center, St. Louis, Missouri, United States of America

Abstract

Heterotrimeric G-proteins comprised of G α , G β and G γ proteins are important signal transducers in all eukaryotes. The G γ protein of the G-protein heterotrimer is crucial for its proper targeting at the plasma membrane and correct functioning. G γ proteins are significantly smaller and more diverse than the G α and G β proteins. In model plants *Arabidopsis* and rice that have a single G α and G β protein, the presence of two canonical G γ proteins provide some diversity to the possible heterotrimeric combinations. Our recent analysis of the latest version of the soybean genome has identified ten G γ proteins which belong to three distinct families based on their C-termini. We amplified the full length cDNAs, analyzed their detailed expression profile by quantitative PCR, assessed their localization and performed yeast-based interaction analysis to evaluate interaction specificity with different G β proteins. Our results show that ten G γ genes are retained in the soybean genome and have interesting expression profiles across different developmental stages. Six of the newly identified proteins belong to two plant-specific G γ protein families. Yeast-based interaction analyses predict some degree of interaction specificity between different G β and G γ proteins. This research thus identifies a highly diverse G-protein network from a plant species. Homologs of these novel proteins have been previously identified as QTLs for grain size and yield in rice.

Citation: Choudhury SR, Bisht NC, Thompson R, Todorov O, Pandey S (2011) Conventional and Novel G γ Protein Families Constitute the Heterotrimeric G-Protein Signaling Network in Soybean. PLoS ONE 6(8): e23361. doi:10.1371/journal.pone.0023361

Editor: Deanna Koepp, University of Minnesota, United States of America

Received: May 19, 2011; **Accepted:** July 13, 2011; **Published:** August 10, 2011

Copyright: © 2011 Choudhury et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: NB was supported by a visiting scholar fellowship from National Institute for Plant Genome Research, India and RT was supported by an National Science Foundation-Research Experiences for Undergraduates grant, (0851838). This work was supported by a United States Department of Agriculture/Agriculture and Food Research Initiative grant to SP, (2010-65116-20454). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: spandey@danforthcenter

^{‡a} Current address: National Institute of Plant Genome Research, New Delhi, India

^{‡b} Current address: University of Hawaii, Hilo, Hawaii, United States of America

Introduction

Heterotrimeric G-proteins comprised of three dissimilar subunits α , β and γ are important signaling intermediates in all eukaryotes [1–3]. The G α subunit, due to its ability to switch between the GDP-bound inactive form and GTP-bound active form, defines the status of signal transduction. Ligand binding to the GPCR causes a change in its conformation allowing an exchange of GTP for GDP on the G α subunit [4]. The GTP-bound G α dissociates from the G $\beta\gamma$ subunits and the released G α -GTP and G $\beta\gamma$ dimer interact with a variety of effector proteins to transduce the signal. The intrinsic GTPase activity of G α causes GTP hydrolysis, converting it back to its GDP-bound state. The G α -GDP reassociates with the G $\beta\gamma$ dimer and the proteins return back to trimeric conformation [4,5].

A wide range of fundamental signal transduction pathways are mediated via G-proteins in both plants and animals [4,6]. In non-plant systems the multiplicity of each of the G-protein subunits, together with almost one thousand GPCRs, tissue-specific expression and signal-dependent heterotrimer formation, provides for the specificity of response [7,8]. In plants the repertoire of G-protein components is relatively simple; the two most investigated species *Arabidopsis* and rice have only a single G α , a single G β and

two canonical G γ proteins [9]. Given the presence of a single G α and G β , the specificity in *Arabidopsis* and rice G-protein signaling is provided solely by the multiplicity of G γ proteins.

We recently carried out an analysis of the soybean genome to assess if this paucity of G-protein components in plants is the norm and whether structural and functional diversity exists within the multiple copies of a gene present in highly duplicated genomes. Our analysis revealed a much more diverse plant G-protein family with the soybean genome encoding for four G α , four G β and two G γ proteins [10]. The number of G α and G β proteins in the soybean genome corresponds well to the two recent genome duplication events [11] resulting in four copies of each gene. Interestingly both G α and G β proteins exhibit some degree of interaction specificity between them. Moreover, based on the GTP-binding and GTPase activity, the four G γ proteins form two distinct subgroups. These data thus revealed that the G-protein signaling in plants is significantly more diverse and complex than what was predicted based on the studies in *Arabidopsis* and rice [10].

The presence of only two G γ proteins in the soybean genome however did not correspond to what was expected based on the genome duplication events. Additionally two of the G β proteins GmG β 1 and GmG β 3 did not show any interaction with the

GmG γ 1 and GmG γ 2 proteins, suggesting that additional G γ proteins may exist. The small size and relatively low sequence conservation make the homology-based identification of G γ proteins difficult; however, they do have certain conserved features to them. All known G γ proteins contain a signature DPLL/I motif which together with few additional conserved amino acids in the middle coiled-coil region is required for interaction with the G β proteins. Most of the known G γ proteins also contain a CAAX motif at C termini which is isoprenylated, resulting in the targeting of the proteins to the plasma membrane [12,13].

In plants G γ proteins have been reported from *Arabidopsis*, rice, pea and soybean. The *Arabidopsis* AGG1 and AGG2 proteins show 48% sequence identity and are involved in regulation of defense responses of plants [14–16]. These two proteins are the prototypical plant G γ proteins. The rice RGG1 and soybean GmG γ 1 and GmG γ 2 proteins are highly homologous to the AGG1 protein and contain all the conserved features and motifs of G γ proteins. The rice RGG2 protein is relatively diverse as this protein has an extra 57 amino acid extension at its N terminus (compared to RGG1) and does not contain the signature prenylation motif. The two reported pea G γ proteins PGG1 and PGG2 are somewhat unusual as they do not contain the highly conserved DPLL/I motif even though a possible prenylation motif is present at its C termini [17]. The function of plant G γ proteins has been evaluated only in *Arabidopsis* where the proteins participate in known G α and/or G β mediated signaling pathways. Molecular-genetic analysis of knockout mutants in AGG1 and AGG2 reveals that the proteins are involved in regulating response to fungal pathogens [16,18].

With the availability of a newer version of the soybean genome assembly (phytozome.net v7) and the use of a series of careful genome annotation programs, we queried the soybean genome to identify additional G γ protein sequences. Our analysis identified two more canonical G γ proteins that are present on the not yet annotated genome regions as well as six additional, novel G γ proteins. The proteins display a great degree of diversity and can be grouped into three distinct families based on sequence features: the archetypal G γ proteins, the prenylation-less G γ proteins and the cysteine-rich G γ proteins. This study describes the identification of these three families of G γ proteins from soybean, details its expression profile in comparison to the expression profile of the GmG β genes and evaluates its interaction with the specific GmG β proteins. The presence of three different families of G γ proteins in a single plant species supports a highly elaborate and diverse G-protein signaling network as well as provides clues to plant-specific G-protein signaling mechanisms, distinct from what is known based on mammalian systems.

Results

Identification of additional canonical and novel G γ proteins from the soybean genome

Our previous analysis of the soybean genome had identified only two G γ proteins [10]. We performed a careful search with the newer version of the soybean genome and identified eight additional G γ proteins. The presence of ten G γ proteins together with four G α and four G β proteins thus corresponds to a total of one hundred and sixty possible heterotrimeric combinations.

Two of the newly identified G γ proteins are highly homologous to the previously identified GmG γ 1 and GmG γ 2 and show high sequence homology between them (Table 1). We named these proteins GmG γ 3 and GmG γ 4, respectively. Both these proteins are present on the regions of the soybean chromosomes that have not yet been annotated. GmG γ 3 is present on the chromosome 20

between the protein coding regions Glyma20g33300.1 and Glyma20g33320.1. We subsequently annotated the locus for GmG γ 3 as Glyma20g33310.1. The open reading frame of this gene along with the positions of introns is reported in Figure S1. GmG γ 4 is present on the chromosome 10 between the protein coding regions Glyma10g32210.1 and Glyma10g32220.1. We annotated the locus for GmG γ 4 as Glyma10g32215.1. Figure S2 details the sequence of this newly annotated gene with its exons and intron.

These four GmG γ proteins G γ 1, 2, 3 and 4 have all the features of canonical G γ proteins, namely the coiled-coil domain in the middle with the conserved DPLL motif at position 66–69 and conserved L30, E40 and S51 (amino acid numbers according to GmG γ 1). These conserved features are important for the interaction of G γ proteins with the G β proteins [19]. GmG γ 3 and GmG γ 4 also contain the CWIL motif at its C termini. This is the most common isoprenylation motif present in all plant G γ proteins. We assigned these four prototypical G γ proteins to group I. Figure 1 shows the protein sequence and conserved motifs of these G γ proteins.

BLAST analysis [20] of the soybean genome using the coiled-coil region of the G γ proteins identified three additional G γ -like proteins on loci Glyma11g18050.1, Glyma14g17060.1 and Glyma17g29590.1. We named these proteins GmG γ 5, GmG γ 6 and GmG γ 7, respectively. These three proteins share a high degree of sequence similarity with each other (Table 1). Based on its unique features, we assigned these three proteins to group II. Compared to the group I proteins, the group II proteins have an extra N terminal extension of 20–25 amino acids. The middle coiled-coil region of these proteins is highly similar to the group I proteins (Figure 1). Sequence features predicted to be involved in G β -G γ interaction are conserved in group II G γ proteins [19]. The most distinct feature of these group II proteins is the lack of C-terminal isoprenylation motif. The proteins end in a RWI sequence instead of a CWIL sequence. The group II proteins are thus somewhat similar to the rice RGG2 proteins as RGG2 also has an N terminal extension (albeit longer, 57 aa) and lacks the prenylation motif [21]. The group II G γ proteins are also small in size: GmG γ 5 and GmG γ 6 are encoded by 131 amino acids and GmG γ 7 is encoded by 126 amino acids, similar to canonical G γ proteins [5]. Surprisingly a prenylation motif-less G γ protein is not present in *Arabidopsis*.

The GmG γ 5 gene sequence is mis-annotated in the current version of the soybean genome. The predicted protein based on the genome annotation is much smaller and does not have the first exon as identified in our study. The genomic arrangement and experimental validation support the presented version of GmG γ 5 as the correct version. The correct sequence of the gene is detailed in Figure S3.

We identified three additional G γ -like novel proteins on loci Glyma15g19630.1, Glyma17g05640.1 and Glyma07g04510.1. We named these proteins GmG γ 8, GmG γ 9 and GmG γ 10, respectively and assigned them to group III based on its distinctive features. The group III proteins are significantly larger than conventional G γ proteins as GmG γ 8, GmG γ 9 and GmG γ 10 are encoded by 213, 228 and 159 amino acids, respectively. The N terminal of group III proteins is similar to group I proteins and the middle coiled-coil domain is highly conserved (Figure 1). GmG γ 8 and GmG γ 9 have all the sequence features required for G β interaction. GmG γ 10 is the only protein identified in our analysis that does not have the conserved DPLL motif, instead this protein contains a similar DPFT motif (Figure 1). It is interesting to note that a highly conserved sequence in mammalian G γ proteins N62 P63, F64 (numbers based on human G γ 1) is not conserved in the

Table 1. Amino acid sequence identity (%) of three different families of soybean G γ proteins.

	GmG γ 1	GmG γ 2	GmG γ 3	GmG γ 4	GmG γ 5	GmG γ 6	GmG γ 7	GmG γ 8	GmG γ 9	GmG γ 10
GmG γ 1	***	92.7	50.9	50.9	29.4	29.4	29.4	20.2	19.3	19.3
GmG γ 2		***	51.9	51.9	31.2	31.2	31.2	19.3	18.3	18.3
GmG γ 3			***	99.1	32.1	32.1	32.1	23.6	24.5	18.9
GmG γ 4				***	32.1	32.1	32.1	23.6	24.5	18.9
GmG γ 5					***	99.2	93.7	20.6	19.8	19.1
GmG γ 6						***	94.4	20.6	19.8	19.1
GmG γ 7							***	21.4	19.0	20.6
GmG γ 8								***	57.3	27.0
GmG γ 9									***	26.4
GmG γ 10										***

doi:10.1371/journal.pone.0023361.t001

plant G γ proteins. This region is required for the GPCR-dependent conformational change in G γ [22].

In addition to large size, there are other features that are unique to group III G γ proteins. The proteins are predicted to have a TNFR (tumor necrosis growth factor receptor) signature and a long cysteine rich C-terminal region which is not found in any other known G γ proteins to date. The C terminus of these proteins is quite variable. Counting from the conserved DPLL (DPFT in case of GmG γ 10) motif, the C terminal of GmG γ 8, GmG γ 9 and

GmG γ 10 consists of 130, 145 and 74 amino acids, respectively. This variable region of group III proteins is unusually high in cysteine content: GmG γ 8 contains 30% Cys (39 out of 130 amino acids), GmG γ 9 contains 33% Cys (49 out of 145 amino acids) and GmG γ 10 contains 26% Cys (19 out of 74 amino acids).

The sequence of *GmG γ 8* and *GmG γ 10* is mis-annotated in the current version of the soybean genome assembly. The predicted genes could never be amplified in our analysis. We manually annotated these genes and amplified the full length product. Based

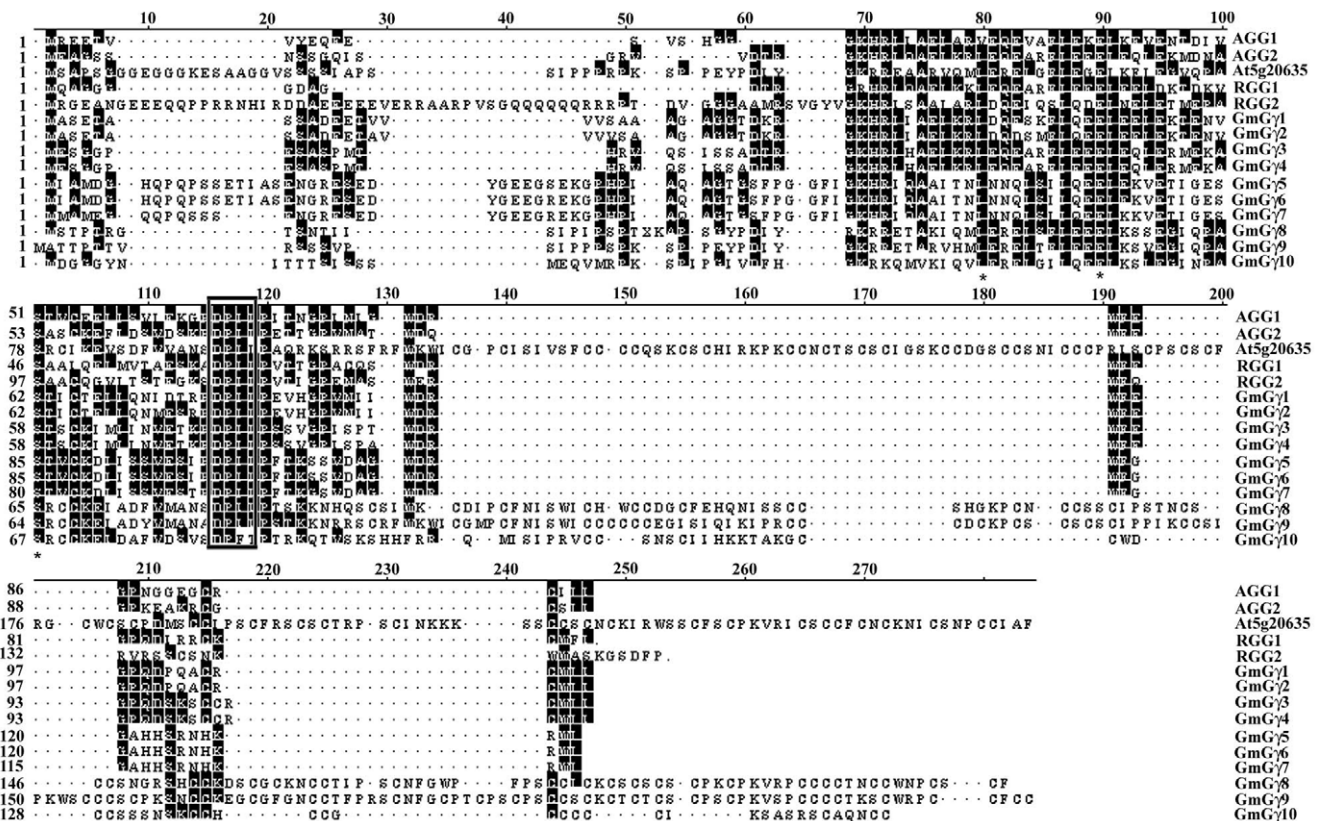


Figure 1. Amino acid sequence alignment of GmG γ proteins. The sequence alignment of *Arabidopsis* AGG1, AGG2 and At5g20635, rice RGG1 and RGG2, and the ten GmG γ was performed using Clustal W (www.clustal.org). Consensus sequences for interaction with G β proteins are marked with asterisks. Box represents the highly conserved DPLL/I motif in G γ proteins. The amino acid positions are numbered in accordance with the GmG γ 1.

doi:10.1371/journal.pone.0023361.g001

on the experimentally obtained cDNAs, we marked the correct exon-intron boundaries of *GmG γ 8* (Figure S4) and *GmG γ 10* (Figure S5).

We also identified a homolog of GmG γ 9 in the *Arabidopsis* genome at locus At5g20635. This gene has been recently described as a G γ protein in *Arabidopsis* [23]. Homologs of group III proteins are present in all plant species. The proteins also show some homology to a keratin associated protein present in mammals. Interestingly the homologs of group III proteins in rice which are named *DEP1* and *GS3* have been recently identified as major QTLs for grain size and yield determination [24,25].

Genome organization and phylogenetic relationship analysis of soybean G γ proteins

The availability of multiple G γ protein sequences in soybean with seemingly variable sequence features raised the question whether these protein families originated from the duplication of a single gene. We analyzed the chromosomal location of all ten *GmG γ* genes and the organization of exon and introns (Figure 2). Group I and group II *GmG γ* genes have four exons each, whereas five exons are present in the group III genes. The length of the second and third exons is highly conserved between all ten *GmG γ* genes. The second exon is 52 bp in group I and II and 53 bp in group III. Similarly, the third exon is 45 bp long in group I and II and 44 bp long in group III *GmG γ* genes. These two exons code for the highly conserved, middle coiled-coil domain of GmG γ proteins. This extreme conservation of the exon organization suggests that the proteins

originated from this core sequence and acquired variable N- and C-terminal sequences. The group I and group II *GmG γ* genes also show a highly conserved fourth exon. It is also interesting to note that despite similarity in the size of cDNAs and proteins of the group I and group II GmG γ s, the group I genes are encoded by large genomic regions (~4–5 kb) and have a very long first intron ranging from 3.5 to 4.5 kb. This feature is not present in the group II *GmG γ* genes. The group III *GmG γ* genes show relatively less conserved genome organization even within the group. The last two exons of this group encode for the cysteine rich region of the proteins and display variability in size (Figure 2).

Phylogenetic analysis of all soybean G γ proteins, along with the *Arabidopsis* and rice sequences, groups them in three expected subgroups with the two *Arabidopsis* proteins, the rice RGG1 and the GmG γ 1-4 in one subgroup, the GmG γ group II proteins and rice RGG2 in another subgroup and the GmG γ group III proteins, the *Arabidopsis* At5g20635 and rice DEP1 and GS3 in the third subgroup (Figure S6). Analysis of chromosome location of *GmG γ* genes suggests that the four group I genes have resulted from two genome duplication events with the first event resulting in two related genes which duplicated again to result in highly homologous *GmG γ 1* and *GmG γ 2* forming one pair and *GmG γ 3* and *GmG γ 4* forming another. Similarly within group II, the genes *GmG γ 5* and *GmG γ 6* form a duplicated gene pair; and within group III, the genes *GmG γ 8* and *GmG γ 9* form a duplicated gene pair. We did not identify duplicated gene pairs corresponding to *GmG γ 7* and *GmG γ 10*. These might have been lost during evolution.

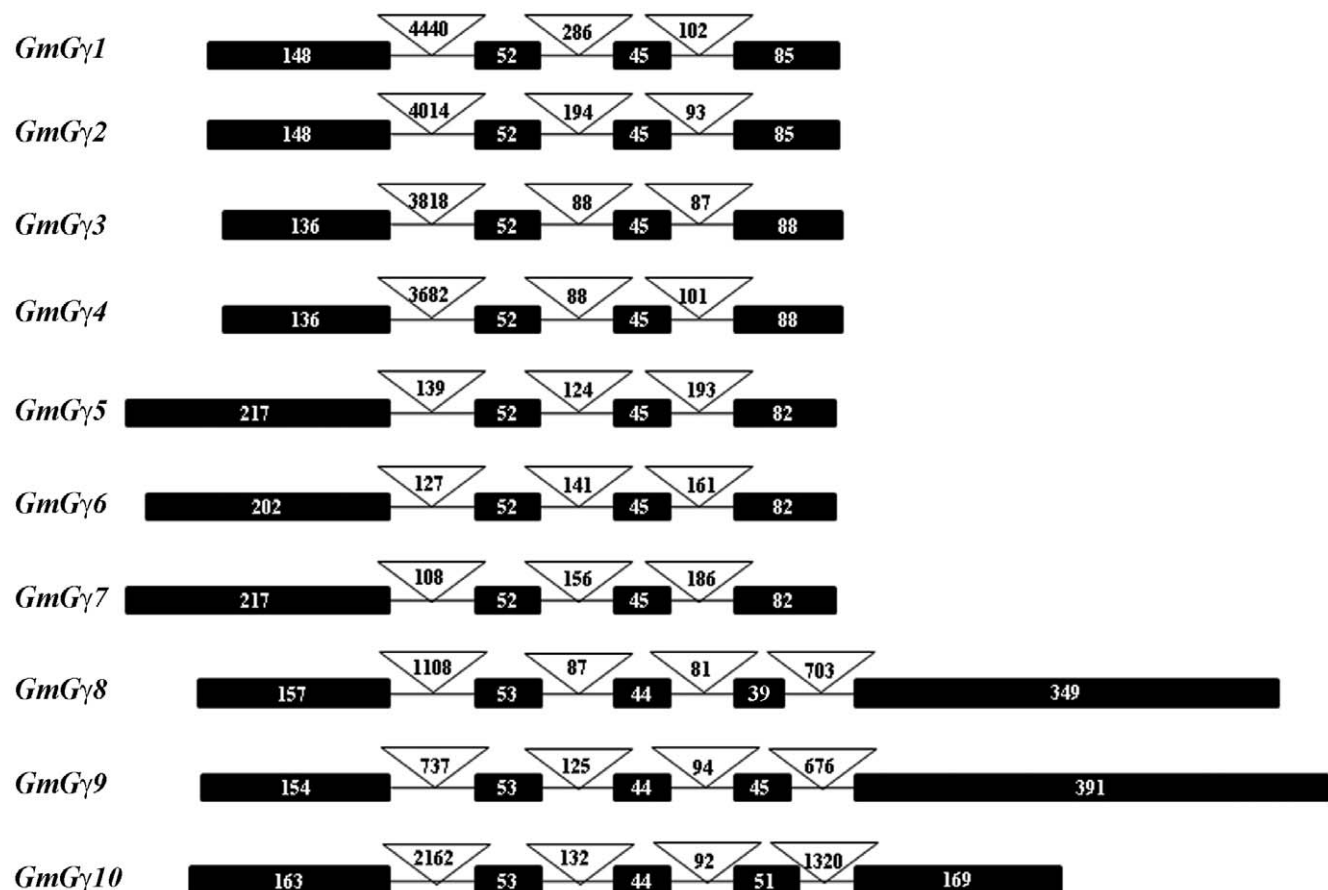


Figure 2. Exon-intron organization of the genomic regions coding for GmG γ proteins. Boxes represent exons and lines represent introns. The introns are not drawn to scale. Numbers denote the length of respective exons and introns.
doi:10.1371/journal.pone.0023361.g002

Tissue-dependent expression analysis of soybean G γ genes

In mammalian systems where multiple isoforms of G γ proteins are present, a high degree of tissue specificity is observed for expression. Similarly the *Arabidopsis* G γ proteins also show distinct expression patterns [16]. We assessed the expression profile of the ten *GmG γ* genes by real-time quantitative PCR to evaluate whether all ten *GmG γ* genes are expressed and the comparative expression levels. We quantified the absolute expression of each gene by a 100 fold serial dilution of cloned plasmid DNA and ascertained the specificity and efficiency of the individual primer pairs (Table S2, Figure S7). A linear correlation coefficient (R^2) of 0.98-0.99 was observed over a 100,000 fold dilution. Interestingly unlike G α and G β genes that are expressed at relatively similar levels in different tissue types [10], a range of variable expression levels were observed for the *GmG γ* genes.

We analyzed the tissue specific expression of the three families of *GmG γ* genes in vegetative tissues and reproductive tissues. Additionally given the possible role of G-protein dependent signaling during nodulation [26–30], we also analyzed the expression of *GmG γ* genes in this legume-specific tissue. Of the group I genes, *GmG γ 4* exhibits overall high expression compared with the other members of this group (Figure 3A). In general, all four genes are expressed in all tissue types tested; however, the expression of *GmG γ 4* is comparatively lower in roots than in aerial tissues whereas *GmG γ 3* is most highly expressed in nodules and at a very low level in developing seeds (S4). The expression data of the genes *GmG γ 1* and *GmG γ 2* are presented for comparing relative expression level to the *GmG γ 3* and *GmG γ 4* genes.

Within the group II genes *GmG γ 6* is expressed at a relatively low level compared to the other two members of this group (Figure 3B). *GmG γ 7* in general has lower expression in leaves at all stages of development. All three genes are expressed at a moderate level in nodules. Additionally the group II genes are expressed at a relatively lower level in reproductive tissues *i.e.* the inflorescence apex, flowers and at an almost non-detectable level in seeds compared to the vegetative tissue (Figure 3B).

The group III genes show maximum variability in tissue specific expression. In this group *GmG γ 9* is expressed at an extremely high level when comparing within this group or with any of the other *GmG γ* genes; whereas *GmG γ 10* is expressed at a very low level. Noticeably this group of genes shows poor expression in nodules and seeds (Figure 3C).

Expression of *GmG γ* genes during seed development and germination

Our previous gene expression analysis with the soybean G α and G β genes showed interesting expression patterns during seed development and germination [10]. We analyzed the expression pattern of the newly discovered *GmG γ* genes during seed development and germination and compared it to the expression of *GmG β* genes.

Specific patterns of gene expression were observed for different *GmG γ* genes. For example *GmG γ 3*, *GmG γ 6* and *GmG γ 7* do not show any change in expression during seed development (Figure 4) similar to *GmG β 1* and *GmG β 4* [10]; whereas *GmG γ 1*, *GmG γ 2* and *GmG γ 9* show moderate down-regulation during seed maturation (Figure 4). Conversely, *GmG γ 4*, *GmG γ 5* and *GmG γ 8* exhibit significant up-regulation of expression during seed maturation (Figure 4). These genes thus show expression profiles similar to *GmG β 2* and *GmG β 3*. It is also noticeable that the expression of these three G γ genes, *GmG γ 4*, *GmG γ 5* and *GmG γ 8*, is maintained at a high level in dry seeds whereas the expression of *GmG β 2* and *GmG β 3* genes returns back to basal levels in dry seeds [10].

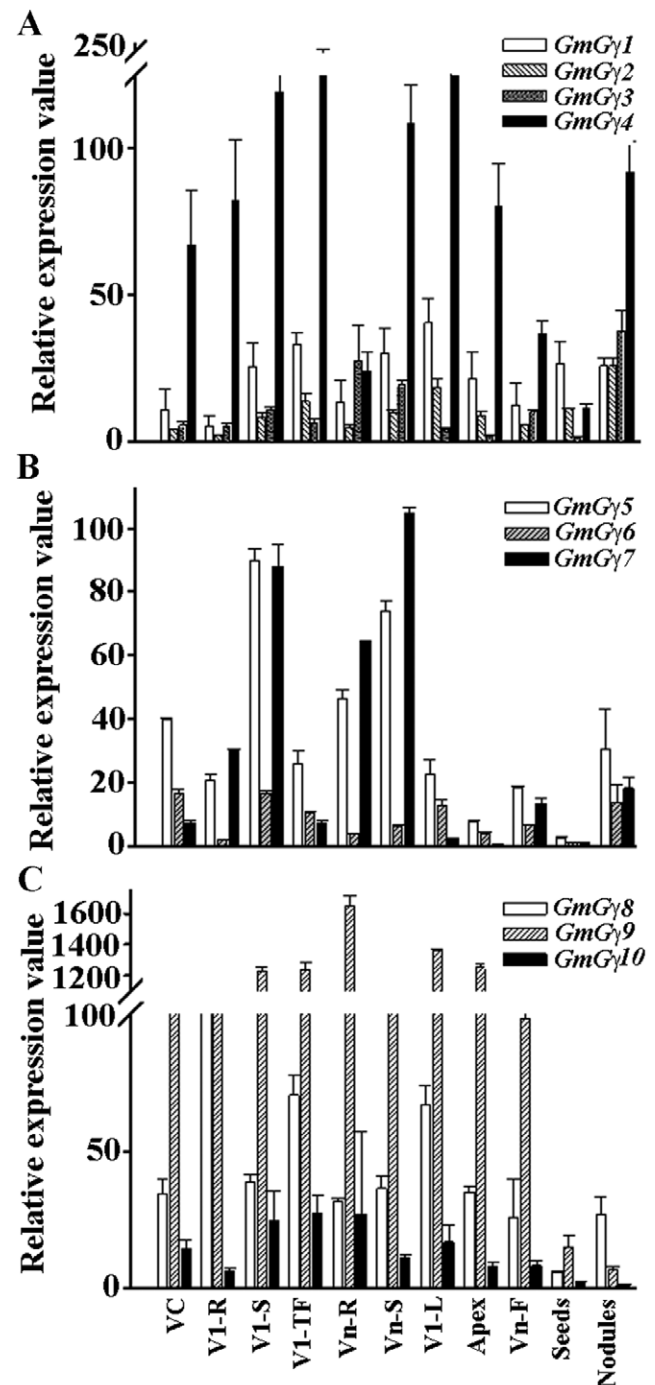


Figure 3. Expression of soybean *GmG γ* genes in different tissue types. The stages are defined as VC: cotyledon; V1-R: primary root at stage V1 (appearance of the first set of unfolded trifoliate leaves); V1-S: primary stem at stage V1; V1-TF: first trifoliate leaf; Vn-R: mature root; Vn-S: mature stem; Vn-L: mature leaves; Apex: inflorescence apex; Vn-F: flower; seeds: developing seeds at stage S4. qRT-PCR amplifications were performed thrice independently for each target, and the data were averaged. The expression values across different tissue types were normalized against soybean *Actin* gene expression [59], which was set at 100. Error bars represent the standard error of the mean. (A) Relative expression of the group I *GmG γ* genes. (B) Relative expression of the group II *GmG γ* genes. (C) Relative expression of the group III *GmG γ* genes.

doi:10.1371/journal.pone.0023361.g003

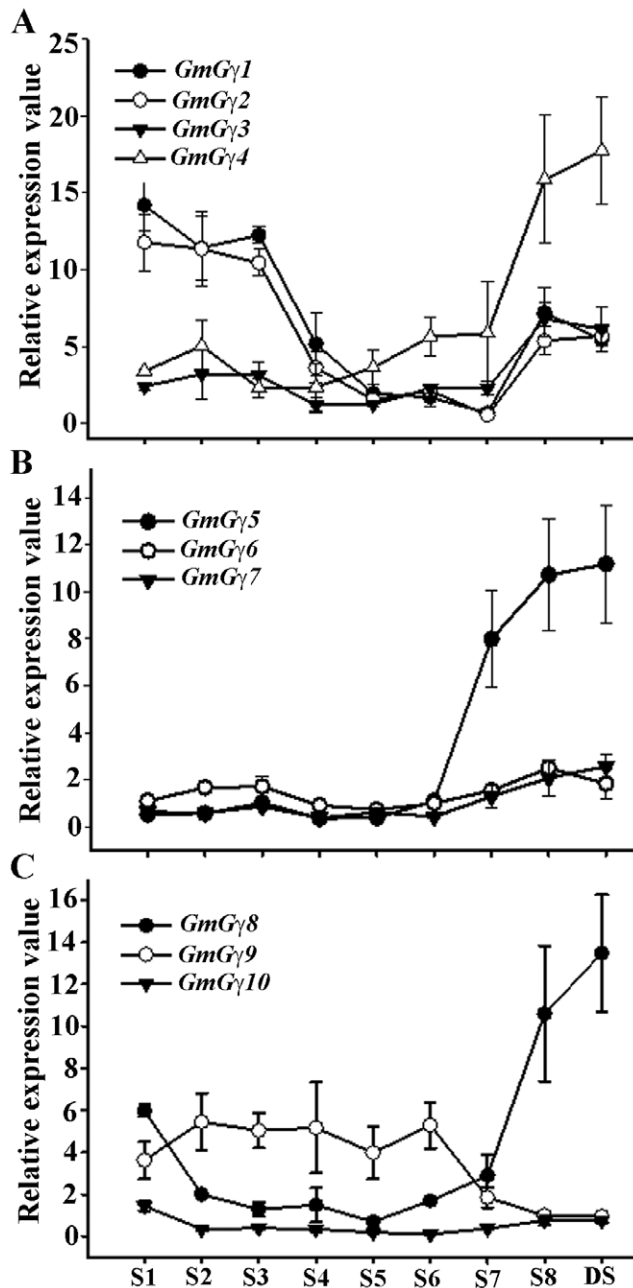


Figure 4. Expression of *GmG γ* genes during different stages of seed development. The seed development stages (S1-S8) are as described in [10]. qRT-PCR amplifications were performed thrice independently for each target, and the data were averaged. The expression values across different seed stages were normalized against soybean *Actin* gene expression. Dry seeds (DS) were also used for the analysis. Error bars represent the standard error of the mean. (A) Relative expression of the group I *GmG γ* genes during different stages of seed development. (B) Relative expression of the group II *GmG γ* genes. (C) Relative expression of the group III *GmG γ* genes. doi:10.1371/journal.pone.0023361.g004

G-proteins are involved in the regulation of seed germination in *Arabidopsis* [31,32] and the soybean *GmG α* and *GmG β* genes show distinct patterns of expression during different stages of seed germination [10]. *GmG γ 3* and *GmG γ 4* follow the similar pattern as *GmG β 3* and *GmG β 4* with higher expression starting 6 h after imbibition, maximizing at 12 h followed by a decrease in

expression (Figure 5). The expression of *GmG γ 5* and *GmG γ 9* is similar to the expression of *GmG β 1* and *GmG β 2* with expression peaking at 6 h after imbibition followed by a decrease in expression (Figure 5). The genes *GmG γ 6* and *GmG γ 10* are expressed at a significantly low level in dry seeds and do not show any change in expression during germination (Figure 5). *GmG γ 7* is also expressed at a very low level in dry seeds (Figure 4C); however, the expression of this gene is significantly up-regulated during germination. The genes show higher expression starting 6 h post-imbibition and maximum expression is observed at 24 h post imbibition followed by a gradual decrease. In contrast *GmG γ 8* which is expressed at a high level in dry seeds shows a significant down-regulation in its expression following 6 h post imbibition.

We also tested the expression of different *GmG γ* genes in response to various stresses. No significant differences were observed in the expression of any of the genes under the conditions where a stress-marker gene *GmRab18* was expressed at significantly higher level (data not shown). These data are similar to what we earlier observed for the *GmG α* and *GmG β* genes [10].

Localization of soybean G γ proteins

Canonical G γ proteins are localized to the plasma membrane via the isoprenylation of C terminal sequence [33,34]. The presence of three G γ protein groups with distinctly variable C terminal sequences allowed us to assess whether the three groups of proteins exhibit any differences in localization. We transiently transformed tobacco leaves via agrobacterium-mediated transformation with YFP (yellow fluorescent protein) fused with respective *GmG γ* genes at N termini. The transformed leaves were visualized with confocal microscopy. The group I fusion proteins YFP-*GmG γ 1*, YFP-*GmG γ 2*, YFP-*GmG γ 3* and YFP-*GmG γ 4* showed fluorescence localized to the periphery of the cells as expected based on the presence of a canonical isoprenylation motif at the C terminal of these proteins (Figure 6). Similarly the group II fusion proteins YFP-*GmG γ 5*, YFP-*GmG γ 6* and YFP-*GmG γ 7* also showed fluorescence restricted to the periphery which suggests the predominantly plasma membrane localization for these proteins (Figure 6). This is intriguing as these proteins lack a canonical prenylation motif and do not have any cysteine residues in the vicinity for such modifications. This is similar to the localization of rice RGG2 protein to the plasma membrane despite lacking a prenylation motif at its C terminal [21].

GmG γ 8 and *GmG γ 9* have a conserved C at position 4 and all three group III genes have a conserved cysteine at position 6 from C-terminal in addition to multiple other cysteine residues in the vicinity. These cysteines qualify for the possible lipid modification by farnesyl transferase or by geranylgeranyl transferases [35]. However, the localization of this group of proteins is interesting as in addition to most of the protein being present at the periphery, fluorescence was also observed as clear puncta especially in the case of YFP-*GmG γ 8* (Figure 6). Such patterns are indicative of either protein aggregate formation or endosomal localizations [34]. Further studies with stably transformed plants will be required to critically assess the localization of group III *GmG γ* proteins.

Protein-protein interaction between *GmG β* and *GmG γ* proteins

We assessed the protein-protein interaction specificity of the three groups of G γ proteins with all four G β proteins of soybean using the ProQuest yeast-2-hybrid system. The group I proteins *GmG γ 3* and *GmG γ 4* showed strong and specific interactions with *GmG β 2* and *GmG β 4* but not with *GmG β 1* and *GmG β 3* proteins

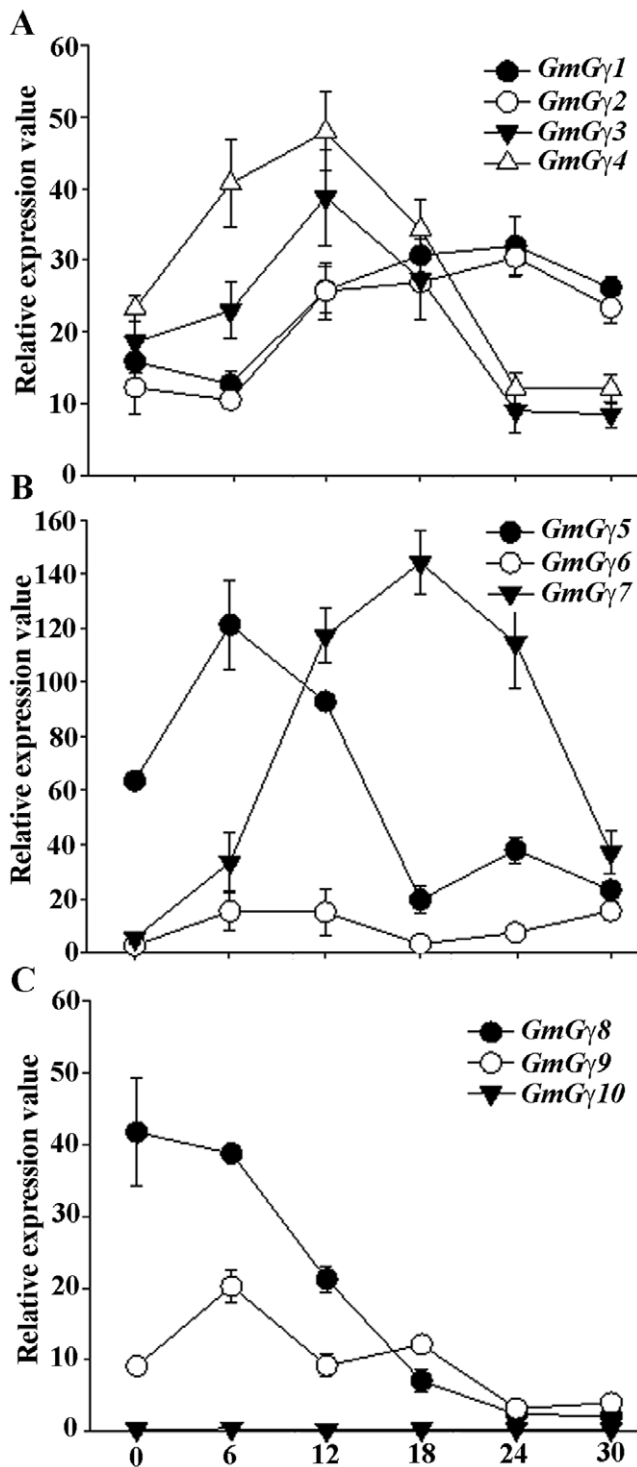


Figure 5. Expression of *GmGγ* genes during seed germination. Seed germination was followed starting from dry seeds (0 h) up to 30 h when an obvious radical had protruded. Seed samples were collected at every 6 h following imbibition. qRT-PCR amplification experiments were performed thrice independently for each target, and the data were averaged. The expression values across different seed germination stages were normalized against soybean *Actin* gene expression. Error bars represent the standard error of the mean. (A) Relative expression of the group I *GmGγ* genes during different stages of seed development. (B) Relative expression of the group II *GmGγ* genes. (C) Relative expression of the group III *GmGγ* genes.
doi:10.1371/journal.pone.0023361.g005

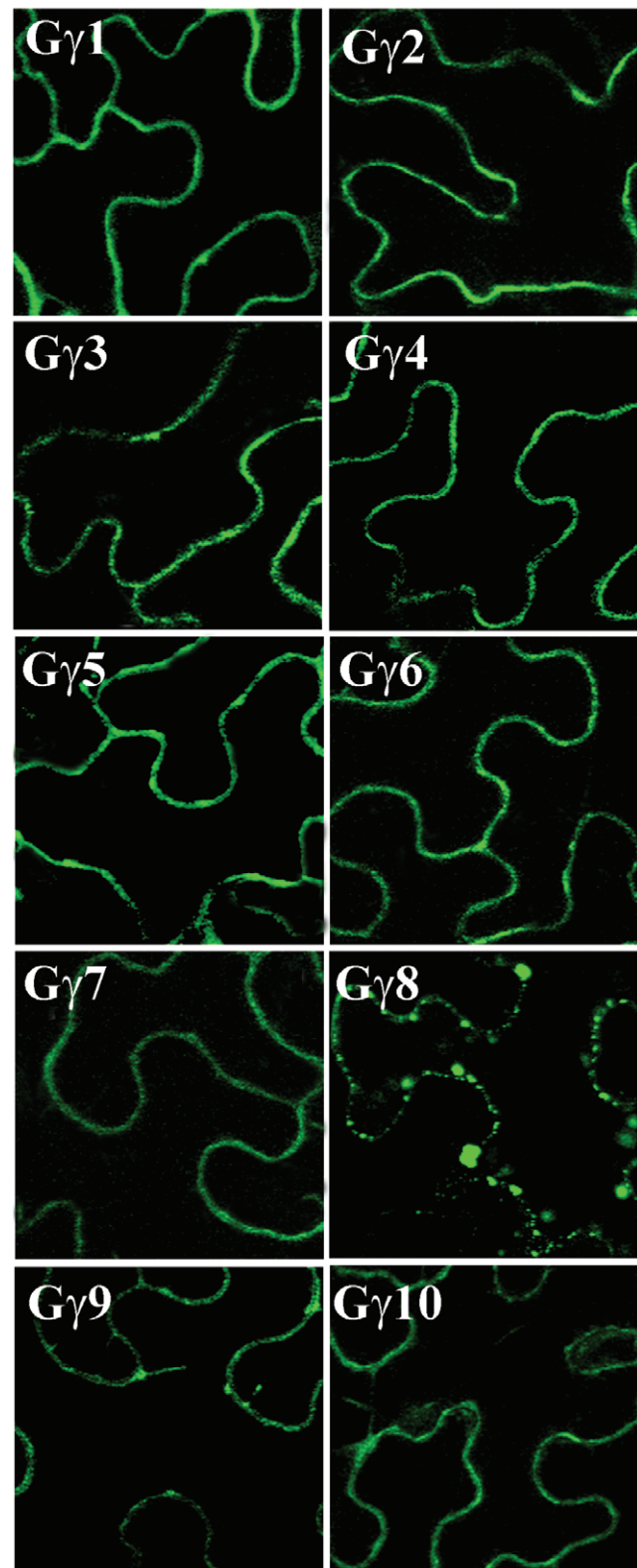


Figure 6. Localization of *GmGγ* proteins. Localization of YFP:*GmGγ* 1-10 genes in transiently transformed tobacco leaves using confocal microscopy. At least three independent transformations were performed. The figure shows representative picture from each transformation.
doi:10.1371/journal.pone.0023361.g006

(Figure 7A) as is also the case with GmG γ 1 and GmG γ 2 [10]. The group II GmG γ proteins exhibit relatively weaker interaction with the GmG β proteins compared to the group I proteins except the interaction between GmG γ 5 and GmG β 4. However, specific interactions of group II proteins were observed with all four GmG β proteins (Figure 7B).

The establishment of interaction with G β proteins was an utmost requirement to classify the group III proteins as novel G γ proteins. As shown in Figure 7C, the group III GmG γ proteins showed very strong and specific interaction with GmG β proteins. Similar to the group I proteins, the group III proteins also interacted strongly with GmG β 2 and GmG β 4. However, distinct from the group I proteins, where no interaction was seen with GmG β 1 and GmG β 3, the group III proteins clearly show weak but specific interaction with GmG β 1 and GmG β 3 proteins (Figure 7C). We also tested whether the N terminus itself was sufficient for this interaction by making deletion constructs of proteins that either expressed the N-terminal half (till DPLL/DPFT motif) or the C-terminal half (protein sequence following the DPLL/DPFT motif). Significantly weaker interaction was detected using either of the truncated proteins (Figure S8), suggesting a full length protein is required for such interactions.

The presence of long cysteine rich C-terminal regions on the group III GmG γ proteins prompted us to test whether these proteins interact with one another. We tested interaction between GmG γ 8, GmG γ 9 and GmG γ 10 in all nine possible combinations. The proteins exhibit strong interaction with each other as shown in Figure 8.

Discussion

Identification of three distinct G γ protein families in soybean

G γ proteins are an integral part of the G-protein heterotrimer. In mammalian systems, these proteins are relatively small (7–8.5 kDa) and are the most diverse of the three subunits with only ~50% sequence homology between different isoforms. The proteins also show a high degree of tissue specific expression and isoform-specific interactions with G β proteins. In both plants and animals, the G γ proteins are required for proper targeting of the G β subunit and of the intact heterotrimer to the plasma membrane [5,36–38]. Identification of diverse G γ proteins from soybean (Figure 1) along with the presence of multiple G α and G β proteins expands the number of possible heterotrimeric combinations in soybean in addition to identifying novel, plant specific components of G-protein signaling.

The C-terminus of G γ proteins is the basis of our classification of the proteins into three distinct groups. Group I is comprised of canonical G γ protein that have all the conserved features of G γ proteins as described based on the mammalian paradigm. Homologs of this family are present in all plant species including gymnosperms and mosses. Most of the reported plant G γ proteins to date, AGG1, AGG2, RGG1 and GmG γ 1–4, are members of this group (Figure 1). The group II proteins consisting of G γ 5, G γ 6 and G γ 7 differ from the family I protein mainly due to the absence of conserved cysteines in the C-terminal region. The genes seem to have evolved from a single amino acid substitution of the CWIL motif to the RWI motif (the most common C-terminal motif present in dicot plants) as most of the protein sequence and exon-intron organization of family I and family II proteins is highly conserved (Figure 2). The rice RGG2 protein is a member of this family although the protein ends in a KGDFS sequence which also seems to be conserved in other monocot species. Lack of cysteine residues in group II proteins precludes the possibility of

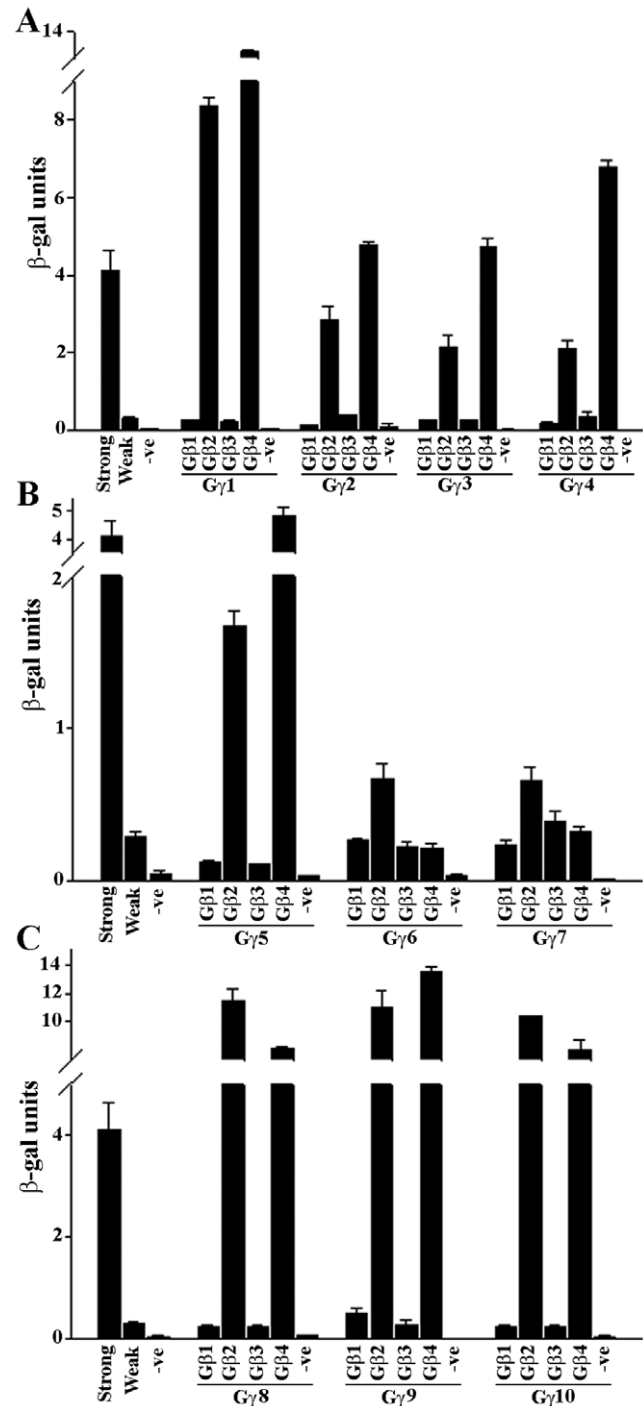


Figure 7. Interaction between soybean G-protein β and γ subunits. Interaction between GmG β (in pDEST32) and GmG γ (in pDEST22) proteins was determined using yeast-2 hybrid-based colorimetric assay. The assays were performed in triplicates and data were averaged. Error bars represent the standard error of the mean. Two biological replicates of the experiment were performed with similar results. (A) Interaction between GmG β proteins and group I GmG γ proteins. (B) Interaction between GmG β proteins and group II GmG γ proteins. (C) Interaction between GmG β proteins and group III GmG γ proteins. Strong, weak and -ve refer to the interaction strength between RalGDS-wt-pDEST32 with Krev1-pDEST22, RalGDS-m1-pDEST32 with Krev1-pDEST22 and RalGDS-m2-pDEST32 with Krev1-pDEST22, respectively. The controls are provided with the ProQuest two hybrid system (Invitrogen). doi:10.1371/journal.pone.0023361.g007

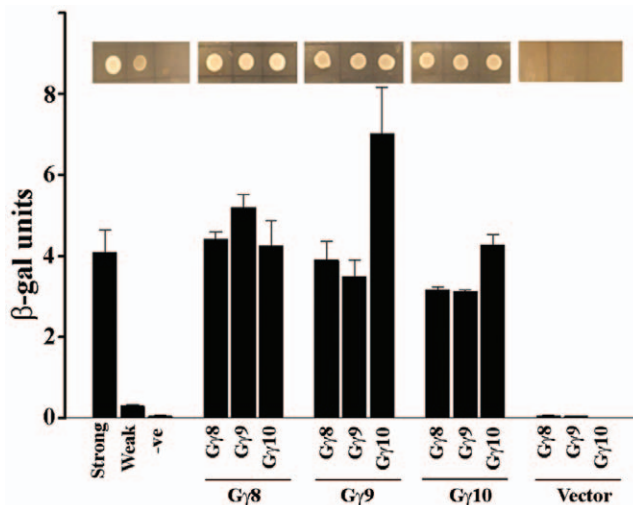


Figure 8. Interaction between group III GmG γ proteins. Interaction between different members of group III GmG γ proteins was determined using yeast-2 hybrid-based growth and colorimetric assay. The assays were performed in triplicates and data were averaged. Error bars represent the standard error of the mean. Two biological replicates of the experiment were performed with similar results. Inset shows growth of yeast colonies on media lacking Leu and Trp but containing 50 μ M 3-AT. Strong, weak and -ve refer to the interaction strength between RalGDS-wt-pDEST32 with Krev1-pDEST22, RalGDS-m1-pDEST32 with Krev1-pDEST22 and RalGDS-m2-pDEST32 with Krev1-pDEST22, respectively.
doi:10.1371/journal.pone.0023361.g008

prenylation; however, the proteins do seem to localize at the plasma membrane (Figure 6) as has also been reported for the rice RGG2 [21]. A single cysteine present in the middle of these proteins could be a potential target for palmitoylation which might assist in its anchoring to the plasma membrane [21]. Additionally the proteins have a high number of positively charged and aromatic amino acids at the C-terminus (7 out of 10) which may target them to the plasma membrane by the formation of an α helix [39]. In mammals, lack of prenylation either by mutation of the conserved cysteine residue or by chemical inhibition has been shown to result in localization of specific G $\beta\gamma$ proteins to the nucleus of the cells and its possible role in regulating transcription, a function not typically associated with G-proteins [40]. Lack of a prenylation-less gene in the *Arabidopsis* genome has limited the functional characterization of this family of G γ in plants. The availability of an insertional mutant line in RGG2 might be able to resolve the issue of whether such proteins play any unconventional roles in plants.

The group III proteins constitute a novel G γ family, specific to plants. Homologs of these proteins are present in both angiosperm and gymnosperm families. We applied the following criteria to establish the group III proteins as authentic G γ proteins. The coiled-coil domain of group III G γ proteins is highly similar to the conventional G γ proteins with full conservation of amino acid residues involved in the interaction with G β proteins (Figure 1). The size of the second and third exons of these proteins is very similar to group I and group II G γ proteins (Figure 2). Additionally a homology modeling-based analysis of three-dimensional protein structures using a fold recognition server Phyre (Protein Homology/analogy Recognition Engine, <http://www.sbg.bio.ic.ac.uk/~phyre/>) predicted these proteins to be G γ proteins with 40–55% precision. Finally the proteins showed strong and specific interaction with the GmG β proteins. The *Arabidopsis* homolog of

group III family protein At5g20635 has recently been identified as a novel G γ protein [23].

Rice has two proteins that show homology to group III proteins, *DEP1* (dense and erect panicle 1) and *GS3* (grain size 3), that have been isolated as major QTLs for seed size and yield [24,25]. Interestingly, the rice G α protein *RGAI* is also involved in regulation of seed size [41]. The rice *DEP1* and *GS3* proteins have been described as novel proteins containing a TNFR motif, a transmembrane domain and proteins with homology to human keratin-associated protein. We also identified a TNFR motif in the soybean group III family proteins; however, using multiple transmembrane prediction programs including TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>), HMMTOP (<http://www.enzim.hu/hmmtop/>) and DAS (<http://www.sbc.su.se/~miklos/DAS/>), we did not identify any transmembrane domains in the GmG γ 8, GmG γ 9 or GmG γ 10 proteins. Rice and *Arabidopsis* group III proteins are predicted to have a single transmembrane domain using DAS, but not with TMHMM or HMMTOP. Experimental verification of the presence of a transmembrane domain and any possible role it might play in localization and/or positioning of these at the plasma membrane will be needed to evaluate its importance. Interestingly, YFP-fused group III GmG γ proteins, in addition to the peripheral YFP fluorescence, also showed small vesicle like structures which were very evident in GmG γ 8. These proteins might be localized to endosomes structures in addition to the plasma membrane. However, since these proteins are highly cysteine rich such structures could also be due to protein aggregate formation or self-interaction (Figure 8). Our data at this time cannot differentiate between these possibilities. Expression of proteins with native promoters in a protein null background will help decipher correct localization.

Expression profile of soybean G γ proteins and possible correlation with G β proteins

The analysis of the complete repertoire of the *GmG γ* genes and its comparison with the expression pattern of *GmG α* and *GmG β* genes began to display specific expression patterns related to particular genes or gene families. Moreover, when comparing the absolute expression levels within different subunits, a wide range of expression levels were observed for the *GmG γ* genes (e.g. Figure 3, *GmG γ 9* versus *GmG γ 6*), whereas all *GmG α* and *GmG β* genes were expressed at a relatively similar level to each other. Additionally the duplicated gene pairs of *GmG α* or *GmG β* typically showed similar expression patterns, a trend not observed between duplicated gene pairs of *GmG γ* genes. *GmG γ 9* was the most highly expressed gene, whereas its duplicated gene *GmG γ 8* was expressed at the moderate level. Likewise, *GmG γ 4* was a highly expressed gene but its duplicated gene pair *GmG γ 3* was relatively poorly expressed.

Some tissue specificity of gene expression was also evident while comparing the expression of multiple G γ genes such as low expression of group II genes in reproductive organs or lower expression of group III genes in nodules (Figure 3). Additionally, during seed development and germination, specific expression profiles were observed for individual genes which in some cases corresponded well to the expression of *GmG β* genes. These observations suggest that developmental stage-specific or tissue-specific expression of particular genes may lead to specific $\beta\gamma$ combinations, similar to what is observed in the mammalian systems [42,43].

Since the two rice homologs of group III genes *DEP1* and *GS3* are involved in grain size determination and yield, we focused on the expression pattern of soybean group III *GmG γ* genes during

seed development. Our data showed that of the three group III genes in soybean, *GmG γ 8* shows the most interesting expression pattern during seed development and germination (Figure 4, 5). The expression of this gene was highly up-regulated when seed is undergoing maturation (stages S7 onwards), whereas a sharp decrease was observed during seed germination. This gene could be a true functional homolog of the rice *DEP1* or *GS3* gene. Additionally the expression of this gene in seeds could be dependent on the endogenous ABA and/or GA concentration as the levels of both these hormones change significantly during seed maturation and germination. Interestingly, *GmG γ 5* also followed a similar expression profile where the expression was up-regulated during seed maturation and generally down-regulated during germination. These genes could be potential targets for manipulation to regulate soybean seed development. It was also obvious that this group of genes is highly variable functionally as *GmG γ 10* shows very little expression in seeds at all developmental stages and exhibits no change in expression profile during germination.

Interaction specificity of GmG β and GmG γ proteins

Specific mammalian G β and G γ proteins form non-dissociable dimers and interact very strongly under a variety of *in vitro* and *in vivo* conditions. The data presented in this study suggest that there is specificity of interaction between different GmG β and GmG γ proteins. It is especially intriguing that the GmG β 1 and GmG β 3 are in general weaker interactors compared to GmG β 2 and GmG β 4 even though they have more than 90% sequence similarity at the protein level (Figure 7). Additionally the group II GmG γ proteins also exhibited weak interactions compared to the group I and group III proteins even though they do interact with similar strengths with all four GmG β proteins. Similar differences in the interaction between mammalian G β and G γ proteins have also been observed. The human G β 1–4 share 80–90% sequence identity; however, G β 1 in general interacts with multiple G γ isoforms, G β 2 is more restricted in its interaction partners and G β 3 displays significantly weaker interactions [44–46]. In most cases, however, the interaction data were based on *in vitro* assays and its relevance in the context of a specific cell type or a signal remains to be evaluated in both mammalian and plant systems.

An intriguing observation in our studies is the strong interaction between different members of the group III proteins themselves (Figure 8). It would be interesting to assess how the oligomerization of these proteins might affect interaction with GmG β proteins or other possible interactors. The unusual nature of these proteins does not preclude the possibility of its involvement in some plant-specific signaling mechanisms which are different from what is known from studies in mammalian systems. Plants do have several unconventional G-protein components such as the extra-large G-proteins that have a G α domain [47–50]; the GTG proteins that have GTP-binding and hydrolysis activity of their own and are regulated by GPA1 [51]; and the RGS1 protein that has a 7TM GPCR-like structure fused to RGS domain [52]. Likewise, most of the known effector proteins of G-protein signaling in plants are also distinct from the conventional effector proteins of mammalian systems. For example, PRN1 (Pirin1) which is a member of an iron-containing subgroup of the cupin superfamily, PD1 (prephenate dehydratase1), a protein involved in phenylalanine biosynthesis, and a NF-Y family transcription factor form a signaling complex during G-protein mediated light and ABA signaling pathways during early growth and development in *Arabidopsis* [53–55]. Similarly, a chloroplast-localized protein THF1 (thylakoid formation 1) is a GPA1 effector protein during sugar signaling [56]. Detailed study of specific pathways mediated

by these unconventional proteins in the context of canonical heterotrimeric G-protein signaling is only in its infancy and future work may divulge additional signaling mechanisms specifically evolved in plants.

Conclusion

We have identified three distinct families of G γ proteins including a novel, plant-specific G γ protein family in the soybean genome. The elucidation of this complete repertoire of different G-protein subunits in soybean reveals a highly elaborate G-protein signaling network in plants. Our data also suggest the presence of subunit-specific and tissue-type or developmental stage-specific heterotrimeric combinations. Additionally the homologs of the group III G γ protein have been identified as major QTLs for grain size and yield in rice. Further work with the generation of RNAi and overexpression lines of soybean G-protein genes will help us decipher its signaling mechanisms as well as its use as potential targets for biotechnological applications.

Materials and Methods

Plant material and growth conditions

Soybean (*Glycine max* L.) cv. Jack seeds were grown in growth chamber (26/20°C day/night temperature, photoperiod of 14/10 h, 800 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity, and 60% humidity). Different developmental stages of soybean plants were collected, immediately frozen in liquid nitrogen and stored at -80°C . Tissue for germination and stress-related experiments was prepared as described in [10].

Cloning of soybean G-protein genes

Soybean G γ genes were identified by analysis of the latest the soybean genome assembly (www.phytozome.net/soybean) with *Arabidopsis* and rice full length and middle coiled-coil region G γ protein sequences as queries. Full-length G γ genes were amplified from soybean seedling cDNA using gene-specific primers (Table S1). The eight newly identified genes were cloned into the pENTR/D-TOPO vector and confirmed by sequencing.

RNA isolation and qRT-PCR

Total RNA was isolated from different tissues of soybean plants using Trizol reagent (Invitrogen) and qRT-PCR experiments were performed essentially according to [10]. The real-time PCR amplification was repeated three times and data were averaged. Sequencing and melt curve analysis of amplicons confirmed specificity. Standard curves for each of the genes were generated using the cloned plasmid DNA of each gene.

Localization of GmG γ proteins

The ten *GmG γ* genes were cloned into the pEarleyGate 104 [57] destination vector using LR clonase mix (Invitrogen). Sequence-confirmed recombinant plasmids containing the *1FP::GmG γ 1–10* were transformed into *A. tumefaciens* strain GV3101 for subsequent plant transformation.

Abaxial surface of tobacco leaves were infiltrated with a log-phase culture of *A. tumefaciens* containing either the gene of interest or an empty vector control according to [58]. Infiltrated plants were incubated in darkness for 36 h followed by 24 h in light. The leaves were imaged with the Zeiss LSM 510 laser scanning confocal microscope (Carl Zeiss) using a 40x water-immersion, 1.2 numerical aperture, C-Apochromat objective. The yellow fluorescent protein (YFP) was excited with the 458-nm line of the argon laser. At least three independent infiltrations were performed for each construct.

Protein-protein interaction assays

To determine the interaction between specific GmG β and GmG γ proteins, GATEWAY-based yeast-two-hybrid assay was performed (ProQuest Two Hybrid System, Invitrogen). Briefly, *GmG β 1-4* genes were cloned into pDEST32 bait vector (containing DNA-binding domain) and *GmG γ 1-10* genes (full length, N-terminal and C terminal parts) were cloned into pDEST22 prey vector (containing DNA-activating domain). The constructs were co-transformed in yeast host strain MaV203 (Invitrogen) in specific combinations. Interaction was determined by growth of diploid yeast colonies on minimal media lacking leucine and tryptophan, but containing 50 μ M 3AT (3-Amino-1,2,4-triazole). The quantitative strength of interaction was determined by β -galactosidase (β -gal) expression assay using ONPG (o-nitrophenyl- β -D-galactopyranoside) as a substrate per manufacturer's instructions.

Supporting Information

Table S1 Gene-specific primers used for expression analysis of *GmG γ* genes.
(DOCX)

Table S2 Absolute copy number quantification and primer amplification efficiency test of *GmG γ* genes.
(DOCX)

Figure S1 Genomic sequence of the newly annotated *GmG γ 3* on chromosome 20.
(DOC)

Figure S2 Genomic sequence of the newly annotated *GmG γ 4* on chromosome 10.
(DOC)

References

- Hamm HE (1998) The many faces of G protein signaling. *J Biol Chem* 273: 669–672.
- Cabrera-Vera TM, Vanhauwe J, Thomas TO, Medkova M, Preinerger A, et al. (2003) Insights into G protein structure, function, and regulation. *Endocr Rev* 24: 765–781.
- Jones AM, Assmann SM (2004) Plants: the latest model system for G-protein research. *EMBO Rep* 5: 572–578.
- Offermanns S (2003) G-proteins as transducers in transmembrane signaling. *Prog Biophys Mol Biol* 83: 101–130.
- McIntire WE (2009) Structural determinants involved in the formation and activation of G protein $\beta\gamma$ dimer. *Neurosignals* 17: 82–99.
- Chen JG (2008) Heterotrimeric G-proteins in plant development. *Front Biosci* 13: 3321–3333.
- Williamson CM, Ball ST, Nottingham WT, Skinner JA, Plagge A, et al. (2004) A cis-acting control region is required exclusively for the tissue-specific imprinting of *Gnas*. *Nat Genet* 36: 894–899.
- Liu J, Chen M, Deng C, Bourc'his D, Nealon JG, et al. (2005) Identification of the control region for tissue-specific imprinting of the stimulatory G protein α -subunit. *Proc Natl Acad Sci U S A* 102: 5513–5518.
- Perfus-Barbeoch L, Jones AM, Assmann SM (2004) Plant heterotrimeric G protein function: insights from Arabidopsis and rice mutants. *Curr Opin Plant Biol* 7: 719–731.
- Bisht NC, Jez JM, Pandey S (2011) An elaborate heterotrimeric G-protein family from soybean expands the diversity of plant G-protein networks. *New Phytol* 190: 35–48.
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, et al. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178–183.
- Fukada Y (1995) Prenylation and carboxylmethylation of G-protein γ subunit. *Methods Enzymol* 250: 91–105.
- Clapham DE, Neer EJ (1997) G protein $\beta\gamma$ subunits. *Annu Rev Pharmacol Toxicol* 37: 167–203.
- Mason MG, Botella JR (2000) Completing the heterotrimer: isolation and characterization of an *Arabidopsis thaliana* G protein γ -subunit cDNA. *Proc Natl Acad Sci U S A* 97: 14784–14788.
- Mason MG, Botella JR (2001) Isolation of a novel G-protein γ -subunit from *Arabidopsis thaliana* and its interaction with G β . *Biochim Biophys Acta* 1520: 147–153.
- Trusov Y, Rookes JE, Tilbrook K, Chakravorty D, Mason MG, et al. (2007) Heterotrimeric G protein γ subunits provide functional selectivity in G $\beta\gamma$ dimer signaling in Arabidopsis. *Plant Cell* 19: 1235–1250.
- Misra S, Wu Y, Venkataraman G, Sopory SK, Tuteja N (2007) Heterotrimeric G-protein complex and G-protein-coupled receptor from a legume *Pisum sativum*: role in salinity and heat stress and cross-talk with phospholipase C. *Plant J* 51: 656–669.
- Trusov Y, Zhang W, Assmann SM, Botella JR (2008) G γ 1+G γ 2 $\neq$ G β : Heterotrimeric G protein G γ -deficient mutants do not recapitulate all phenotypes of G β -deficient mutants. *Plant Physiol* 147: 636–649.
- Temple BR, Jones AM (2007) The plant heterotrimeric G-protein complex. *Annu Rev Plant Biol* 58: 249–266.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.
- Kato C, Mizutani T, Tamaki H, Kumagai H, Kamiya T, et al. (2004) Characterization of heterotrimeric G protein complexes in rice plasma membrane. *Plant J* 38: 320–331.
- Gautam N (2003) A conformational switch regulates receptor-G protein interaction. *Structure* 11: 359–360.
- Chakravorty D, Trusov Y, Zhang W, Acharya BR, Sheahan MB, et al. (2011) An atypical heterotrimeric G protein γ subunit is involved in guard cell K (+) channel regulation and morphological development in *Arabidopsis thaliana*. *Plant J* (in press);doi: 10.1111/j.1365-3113.2011.04638.x.
- Huang X, Qian Q, Liu Z, Sun H, He S, et al. (2009) Natural variation at the DEP1 locus enhances grain yield in rice. *Nat Genet* 41: 494–497.
- Mao H, Sun S, Yao J, Wang C, Yu S, et al. (2010) Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc Natl Acad Sci U S A* 107: 19579–19584.
- Pingret JL, Journet EP, Barker DG (1998) Rhizobium nod factor signaling: Evidence for a G protein-mediated transduction mechanism. *Plant Cell* 10: 659–672.
- den Hartog M, Musgrave A, Munnik T (2001) Nod factor-induced phosphatidic acid and diacylglycerol pyrophosphate formation: a role for phospholipase C and D in root hair deformation. *Plant J* 25: 55–65.
- Kelly MN, Irving HR (2003) Nod factors activate both heterotrimeric and monomeric G-proteins in *Vigna unguiculata* L Walp. *Planta* 216: 674–685.
- Sun J, Miwa H, Downie JA, Oldroyd GE (2007) Mastoparan activates calcium spiking analogous to Nod factor-induced responses in *Medicago truncatula* root hair cells. *Plant Physiol* 144: 695–702.
- Santos-Briones CDL, Cárdenas L, Estrada-Navarrete G, Santana O, Minero-García Y, et al. (2009) GTP γ S antagonizes the mastoparan-induced in vitro activity of PIP $_2$ -phospholipase C from symbiotic root nodules of *Phaseolus vulgaris*. *Physiol Plantarum* 135: 237–245.

Figure S3 Correct genomic sequence of *GmG γ 5* as experimentally verified.
(DOC)

Figure S4 Correct genomic sequence of *GmG γ 8* as experimentally verified.
(DOC)

Figure S5 Correct genomic sequence of *GmG γ 10* as experimentally verified.
(DOC)

Figure S6 Evolutionary relationships of GmG γ proteins.
(PPT)

Figure S7 PCR efficiency of *GmG γ* genes over 100,000 fold dilution.
(PPT)

Figure S8 Test of interaction between the N-terminal and C-terminal parts of GmG γ 8, 9 and 10 proteins with different GmG β proteins.
(PPT)

Acknowledgments

We thank Drs. Jan A. Meirnyk, Christopher Taylor, Senthil Subramanian and Manjula Govindarajulu for their help with various soybean-related resources and Ms. Christine Ehret for critical editing of the manuscript.

Author Contributions

Conceived and designed the experiments: SC NB SP. Performed the experiments: SC NB RT OT SP. Analyzed the data: SC NB SP. Wrote the paper: SP.

31. Ullah H, Chen JG, Wang S, Jones AM (2002) Role of a heterotrimeric G protein in regulation of Arabidopsis seed germination. *Plant Physiol* 129: 897–907.
32. Pandey S, Chen JG, Jones AM, Assmann SM (2006) G-protein complex mutants are hypersensitive to abscisic acid regulation of germination and postgermination development. *Plant Physiol* 141: 243–256.
33. Simonds WF, Butrynski JE, Gautam N, Unson CG, Spiegel AM (1991) G-protein $\beta\gamma$ dimers. Membrane targeting requires subunit coexpression and intact γ C-A-A-X domain. *J Biol Chem* 266: 5363–5366.
34. Takida S, Wedegaertner PB (2003) Heterotrimer formation, together with isoprenylation, is required for plasma membrane targeting of G $\beta\gamma$. *J Biol Chem* 278: 17284–17290.
35. Crowell DN, Huizinga DH (2009) Protein isoprenylation: the fat of the matter. *Trends Plant Sci* 14: 163–170.
36. Adjobo-Hermans MJ, Goedhart J, Gadella TW (2006) Plant G protein heterotrimers require dual lipidation motifs of G α and G γ and do not dissociate upon activation. *J Cell Sci* 119: 5087–5097.
37. Zeng Q, Wang X, Running MP (2007) Dual lipid modification of Arabidopsis G γ -subunits is required for efficient plasma membrane targeting. *Plant Physiol* 143: 1119–1131.
38. Smrcka AV (2008) G protein $\beta\gamma$ subunits: central mediators of G protein-coupled receptor signaling. *Cell Mol Life Sci* 65: 2191–2214.
39. Prinz WA, Hinshaw JE (2009) Membrane-bending proteins. *Crit Rev Biochem Mol Biol* 44: 278–291.
40. Kino T, Kozasa T, Chrousos GP (2005) Statin-induced blockade of prenylation alters nucleocytoplasmic shuttling of GTP-binding proteins $\gamma 2$ and $\beta 2$ and enhances their suppressive effect on glucocorticoid receptor transcriptional activity. *Eur J Clin Invest* 35: 508–513.
41. Ueguchi-Tanaka M, Fujisawa Y, Kobayashi M, Ashikari M, Iwasaki Y, et al. (2000) Rice dwarf mutant d1, which is defective in the α subunit of the heterotrimeric G protein, affects gibberellin signal transduction. *Proc Natl Acad Sci U S A* 97: 11638–11643.
42. Wilcox MD, Dingus J, Balcueva EA, McIntire WE, Mehta ND, et al. (1995) Bovine brain GO isoforms have distinct gamma subunit compositions. *J Biol Chem* 270: 4189–4192.
43. Oldham WM, Hamm HE (2008) Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat Rev Mol Cell Biol* 9: 60–71.
44. Pronin AN, Gautam N (1992) Interaction between G-protein β and γ subunit types is selective. *Proc Natl Acad Sci U S A* 89: 6220–6224.
45. Schmidt CJ, Thomas TC, Levine MA, Neer EJ (1992) Specificity of G protein β and γ subunit interactions. *J Biol Chem* 267: 13807–13810.
46. Dingus J, Wells CA, Campbell L, Cleator JH, Robinson K, et al. (2005) G Protein $\beta\gamma$ dimer formation: G β and G γ differentially determine efficiency of in vitro dimer formation. *Biochemistry* 44: 11882–11890.
47. Assmann SM (2002) Heterotrimeric and unconventional GTP binding proteins in plant cell signaling. *Plant Cell* 14: S355–S373.
48. Lee YR, Assmann SM (1999) Arabidopsis thaliana ‘extra-large GTP-binding protein’ (AtXLG1): a new class of G-protein. *Plant Mol Biol* 40: 55–64.
49. Ding L, Pandey S, Assmann SM (2008) Arabidopsis extra-large G proteins (XLGs) regulate root morphogenesis. *Plant J* 53: 248–263.
50. Pandey S, Monshausen GB, Ding L, Assmann SM (2008) Regulation of root-wave response by extra large and conventional G proteins in *Arabidopsis thaliana*. *Plant J* 55: 311–322.
51. Pandey S, Nelson D, Assmann SM (2009) Two novel GPCR-type G-proteins are abscisic acid receptors in Arabidopsis. *Cell* 136: 136–148.
52. Chen JG, Willard FS, Huang J, Liang J, Chasse SA, et al. (2003) A seven-transmembrane RGS protein that modulates plant cell proliferation. *Science* 301: 1728–1731.
53. Warpeha KM, Lateef SS, Lapik Y, Anderson M, Lee BS, et al. (2006) G-protein-coupled receptor 1, G-protein Galpha-subunit 1, and prephenate dehydratase 1 are required for blue light-induced production of phenylalanine in etiolated Arabidopsis. *Plant Physiol* 140: 844–855.
54. Warpeha KM, Upadhyay S, Yeh J, Adamiak J, Hawkins SI, et al. (2007) The GCR1, GPA1, PRN1, NF-Y signal chain mediates both blue light and abscisic acid responses in Arabidopsis. *Plant Physiol* 143: 1590–1600.
55. Warpeha KM, Gibbons J, Carol A, Slusser J, Tree R, et al. (2008) Adequate phenylalanine synthesis mediated by G protein is critical for protection from UV radiation damage in young etiolated *Arabidopsis thaliana* seedlings. *Plant Cell Environ* 31: 1756–1770.
56. Huang J, Taylor JP, Chen JG, Uhrig JF, Schnell DJ, et al. (2006) The plastid protein THYLAKOID FORMATION1 and the plasma membrane G-protein GPA1 interact in a novel sugar-signaling mechanism in Arabidopsis. *Plant Cell* 18: 1226–1238.
57. Earley KW, Haag JR, Pontes O, Oppen K, Juehne T, et al. (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J* 45: 616–629.
58. Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J* 33: 949–956.
59. Subramanian S, Xu L, Lu G, Odell J, Yu O (2004) The promoters of the isoflavone synthase genes respond differentially to nodulation and defense signals in transgenic soybean roots. *Plant Mol Biol* 54: 623–639.