

RESEARCH ARTICLE

# Rule-Based Design of Plant Expression Vectors Using GenoCAD

Anna Coll<sup>1\*</sup>, Mandy L. Wilson<sup>2</sup>, Kristina Gruden<sup>1</sup>, Jean Peccoud<sup>2</sup>

**1** Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia, **2** Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, Virginia, United States of America

\* [anna.coll@nib.si](mailto:anna.coll@nib.si)



**OPEN ACCESS**

**Citation:** Coll A, Wilson ML, Gruden K, Peccoud J (2015) Rule-Based Design of Plant Expression Vectors Using GenoCAD. PLoS ONE 10(7): e0132502. doi:10.1371/journal.pone.0132502

**Editor:** Martina Stromvik, McGill University, CANADA

**Received:** March 8, 2015

**Accepted:** June 15, 2015

**Published:** July 6, 2015

**Copyright:** © 2015 Coll et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files and also available from Figshare (<http://dx.doi.org/10.6084/m9.figshare.1428589>; <http://dx.doi.org/10.6084/m9.figshare.1327969>; <http://dx.doi.org/10.6084/m9.figshare.1327970>).

**Funding:** This work is supported by National Science Foundation Awards 1241328, JP and The Slovenian Research Agency Program P4-0165 and Project N4-0026, KG.

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

## Abstract

Plant synthetic biology requires software tools to assist on the design of complex multi-genic expression plasmids. Here a vector design strategy to express genes in plants is formalized and implemented as a grammar in GenoCAD, a Computer-Aided Design software for synthetic biology. It includes a library of plant biological parts organized in structural categories and a set of rules describing how to assemble these parts into large constructs. Rules developed here are organized and divided into three main subsections according to the aim of the final construct: protein localization studies, promoter analysis and protein-protein interaction experiments. The GenoCAD plant grammar guides the user through the design while allowing users to customize vectors according to their needs. Therefore the plant grammar implemented in GenoCAD will help plant biologists take advantage of methods from synthetic biology to design expression vectors supporting their research projects.

## Introduction

Synthetic biology aims at bioengineering organisms that perform beneficial functions, generally by means of a rational design approach [1]. Plants have largely been unexploited for synthetic biology, but they offer great potential [2]. To fully benefit from synthetic biology, significant efforts have been dedicated to the development of robust, less-demanding, and more reliable methods to assemble increasingly complex designs (see review [3]). Beyond the assembly of constructs, the design of complex multigene vectors is a big challenge. Editing large DNA sequences increases the risk of introducing errors. Furthermore, identifying suitable biological parts is becoming more difficult as the number of parts for synthetic biology increases. Therefore, there is a need for software tools that guide plant synthetic biologists through the design of application-specific expression vectors. GenoCAD ([www.genocad.com](http://www.genocad.com)) is a Computer-Aided Design (CAD) software for synthetic biology which allows the user to quickly design protein expression vectors, artificial gene networks and other genetic constructs based on the notion of genetic parts [4]. GenoCAD includes a system to manage annotated and user-defined genetic parts. Moreover, it also guides the users through the design by means of a set of predefined rules that describe the design strategy for a specific type of application and which can be expressed in a context-free grammar [5]. By default, GenoCAD includes a simple

grammar used for demonstration purposes. However, the grammar editor embedded in GenoCAD enables users to develop brand new grammars, and therefore provides biologists with a tool to formalize custom design strategies. Several grammars have already been added by users i.e. (i) a grammar to design a family of vaccine vectors derived from vesicular stomatitis virus (VSV) [6], (ii) a grammar to design *Chlamydomonas reinhardtii* expression vectors [7] and (iii) a grammar to design synthetic transcription factors in eukaryotes [8].

Here, we present a vector design strategy to express genes in plants. Specifically, we developed a grammar to design constructs for three categories of experiments: promoter analysis, protein localization and protein-protein interaction (PPI) studies, which are perhaps the most demanding in plant biology. The design rules of these vectors are sufficiently different from the design of protein expression in *E.coli* to justify the development of a new grammar following the grammar-design workflow previously described by Wilson et al [7]. Constructs for promoter analysis, localization, and PPI studies are frequently used in the plant biology community. As can be seen in the literature they are the methods of choice to functionally characterize plant proteins. For instance, among other examples, these methodologies were used to elucidate the role of MAPKKKε in plant immune response against *Phytophthora infestans* [9], to study the function of ERF3 in potato response to various biotic and abiotic stresses [10] or to understand the role of *Arabidopsis* WRKY8 in mediating salt stress tolerance [11]. Many commercially available plasmids can be used for these applications, but they do not always meet all the requirements of specific projects. This usually forces the user to adapt the vector, and edit the sequence using a cut-and-paste approach, which has a high chance of introducing errors. Here we have captured the expertise gained by our group over the years to produce set of design rules and a parts library, which will guide non-expert users through the design process. Advanced users can use the GenoCAD grammar editor to customize this grammar by adding rules and parts to extend the scope of the grammar. Alternatively, they could also simplify the grammar by eliminating specific rules that do not apply to their projects, or even delete an entire branch of the grammar to streamline the design process.

## Methods

### Category definition

First, we identified different categories of genetic parts used to design plant expression vectors (S1 File). Our plant grammar includes several general categories found in many other grammars e.g. gene (GEN), epitope tag (ETG), fluorescent tag (FTG), linkers (LNK), terminator (TER) or promoter (PRO) (S1 Table). The category PRO refers to the most commonly used promoters for expression in plants. These are constitutive and strong promoters usually of viral origin. In this plant grammar we distinguished them from native promoters (NPRO), which includes promoters naturally found in plants that regulate the expression of plant genes. Like for the *C. reinhardtii* chloroplast grammar [7], start (ATG) and stop (STP) codon are considered as categories in order to facilitate the design of fusion proteins. Since 5'UTR and 3'UTR are not well characterized in plants, here they are included in the promoter and terminator category respectively.

In addition to terminal categories corresponding to groups of genetic parts, a grammar generally includes rewritable categories corresponding to higher levels of organization. For instance, we defined the categories 'one/two promoter' (PRO12) and 'one/two terminator' (TER12) to allow the option of using a single or double promoter and terminator. We also defined categories specific to each application (S1 Table). For instance, GFTG is a rewritable category consisting of a gene fused to a FTG. It is a specific category for protein localization studies. These categories will be explained in detail in Results.

The grammar also includes categories used as delimiters. Parentheses are used to delimit the sequence of a plasmid. They can be used to indicate that a design corresponds to a complete plasmid. They can also be used in situations where an experiment calls for a co-design of two plasmids working together.

The brackets [] are used to indicate the orientation of a sequence. Elements between brackets are coded on the negative strand. DNA sequences of elements between brackets are reverse-complemented when exported as a text file.

## Rules

Rules for the plant grammar ([S1 File](#)) are shown in [S2 Table](#). S is the start symbol, and it can be rewritten into an expression plasmid for localization studies (*loc* route), for promoter analysis (*pro* route) or into plasmids for PPI studies (*ppi* route). There are several common rules for *ppi* and *loc* routes to rewrite the GEN, PRO and TER categories. There are 5 rules to rewrite the GEN category. Rule *2gen* is used to fuse two proteins. Rules *tgen* and *gent* allow expressing a protein fused to an ETG on its N or C terminus. Epitope tags are widely used for detection or purification purposes. The rules *lgen* and *genl* are used to include a LNK between fusion proteins. Linkers are routinely used when expressing recombinant fusion proteins in order to improve expression, folding and stability of the proteins [12]. Three rules can be used to rewrite the PRO12 category; rule *pro1* is used to add a single promoter, rule *pro2* allows the user to introduce an additional promoter, and rule *npro* introduces a native promoter. Similarly, the TER12 category can be rewritten into either a single or double terminator. This rule is also applicable to the route for promoter analysis. Moreover, *pro* and *loc* routes share rules to rewrite the FTG category. The fluorescent tag can be rewritten in both cases as a protein with epitope tags and/or linker domains at its N and/or C terminus.

Specific rules for each route are described in the Results section.

## Parts Library

The library was built by importing sequences into each category ([S1 File](#)). The sources of part sequences are available in the parts description field. Under GEN and NPRO categories we imported sequences from *S. tuberosum* group Phureja DM1-3 [13] as an example since these are the sequences experimentally tested in our laboratory.

The use of FTGs in plant biology is especially challenging because of the autofluorescence displayed by several components of plant tissues (chlorophyll, lignified secondary cell walls, and vacuolar contents) overlap with the emission wavelength of green fluorescent protein (GFP) [14]. The FTGs currently added in the GenoCAD parts library have all been previously tested in *Nicotiana benthamiana* and *Solanum tuberosum* leaves, and they include enhanced GFP (eGFP), yellow fluorescent protein (eYFP), cyan fluorescent protein (eCFP) and mCherry.

Several combinations of fluorescent protein fragments have been functionally tested and recommended for bimolecular fluorescence complementation (BiFC) analysis [15]. All of them present advantages and drawbacks in terms of complementation efficiency and fluorescence intensity. Fragments of YFP truncated at residue 155 [16] were incorporated in the parts library because, according to Kerppola et al. [17] and our own experience, though having weak fluorescent intensity, this combination of FTG fragments exhibits low spontaneous association, and, therefore, less possibility of false positive results than others.

Vector backbones to clone the assembled parts currently available in the library include four different pCAMBIA minimal selection vectors (<http://www.cambia.org>). These vectors are T-DNA binary vectors (see review [18]), and are therefore compatible with Agrobacteria-mediated plant transformation. They contain minimal heterologous sequences for plant

transformation and selection of transformants (in bacteria and plant); they allow for the insertion of desired genes for transformation into plants, but require all regulatory sequences for plant expression of newly cloned genes. All vectors were opened at the multiple cloning site (MCS) with *Sall* and *BamHI* restriction enzymes.

## Results

Plant transformation for functional analysis experiments has become a routine tool in plant research. Nowadays, commercial plant expression plasmids for different applications are available. However, this is a rigid option, and to change any element of the backbone plasmid is a time-consuming task. Moreover, most of them are based on classical or Gateway cloning systems. There is a need for simple and versatile design strategies to allow high throughput approaches in synthetic biology studies. Therefore, we developed a GenoCAD grammar to design constructs for *in planta* functional analysis studies ([S1 File](#)). The plant grammar is available in Figshare ([doi.org/10.6084/m9.figshare.1428589](https://doi.org/10.6084/m9.figshare.1428589)).

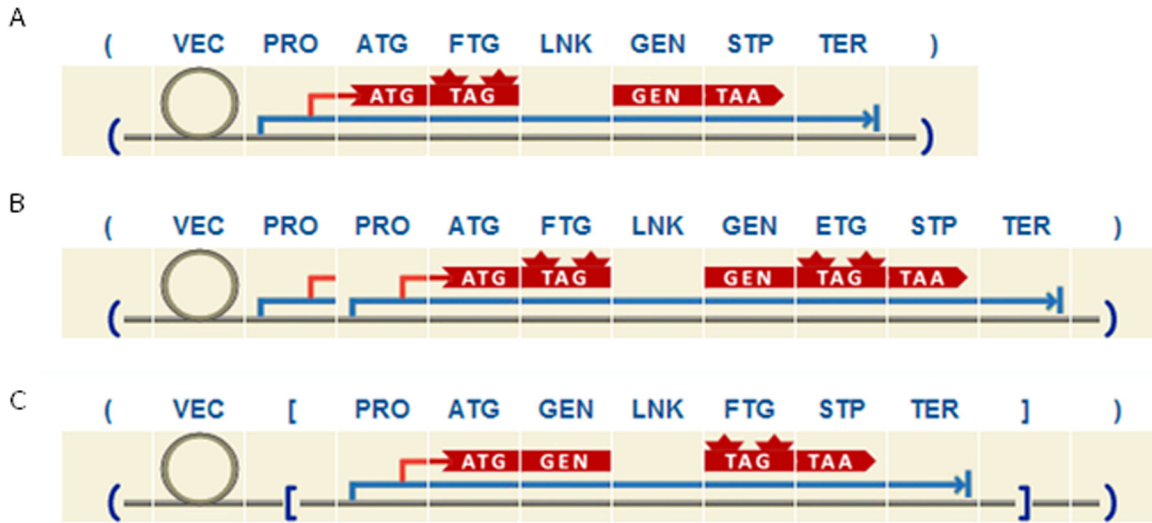
### Design of expression plasmids for protein localization studies

Studying the protein subcellular localization in living plant cells is a useful tool for characterization of unknown proteins. Fusion of fluorescent tags to proteins of interest has become the method of choice for this purpose [19]. The availability of FTG markers has increased noticeably over the last two decades. The large selection of FTG markers available in plant molecular biology together with the development of sophisticated image acquisition and analysis software contributed to the rapidly and increased adoption of this technology by plant molecular biologists [14]. Therefore, we implemented a grammar to design constructs for protein localization studies based on fused FTG.

Route *loc* guides the user through the design of plant expression vectors for localization studies ([S2 Table](#)). By means of *lcas* rule, localization category (LOC) is rewritten into a complete plasmid that includes an expression cassette (CAS) and vector backbone (VEC). The next rule is *prct*, which breaks the cassette down into PRO12, open reading frame (CDS) and TER12. The orientation of the expression cassette can be reversed by using *rcas* rule. Rules *npro*, *pro1*, *pro2*, *ter1* and *ter2* allow the user to design an expression cassette with a native promoter, single or double promoter, and single or double terminator. Afterwards, rule *gnftg* break the CDS down into an ATG, GFTG, and STOP codon. Therefore, this rule constrains the user to design an expression cassette with a gene fused at least with one fluorescent protein, which is the minimal requirement for plasmids with localization application purposes. Two rules (*nftg* and *cftg*) incorporate flexibility into the design, allowing users to add the fluorescent protein on the 5' or 3'-end of the gene. Moreover, epitope tags and/or linker domains can also be added at both sides of the FTG (rules *tftg*, *ftgt*, *lftg* and *ftgl*). Lastly, as it was described, GEN category can be rewritten in order to add other open reading frames, epitope tags and/or linkers at the 5' and/or 3'-end of the gene. To demonstrate the flexibility of this approach, 3 examples of different designs with different degrees of complexity, all of them applicable for protein localization studies, are shown in [Fig 1](#).

### Design of promoter analysis studies

Identification and characterization of plant promoters is crucial to understanding the function of the genes under their control. Characterizing the promoters themselves is also important because they are valuable tools in plant genetic engineering (see [20] as an example). Therefore, the promoter-reporter analysis is an indispensable approach for molecular plant biologists. In



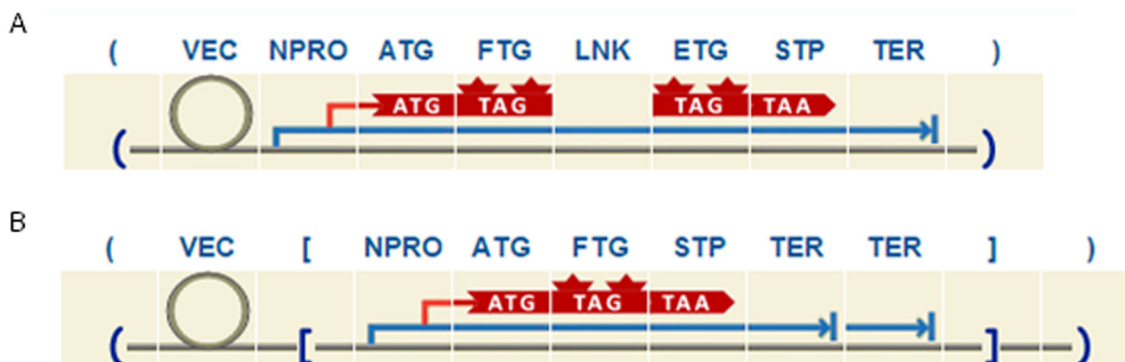
**Fig 1. Example of three different designs for localization studies purposes as developed with the plant grammar.** A. Scheme of the most basic structure we can design, where the expression cassette includes the GEN fused to a FTG by means of a LNK domain on the N terminal. B. Sample design includes an expression cassette with 2 PRO and a GEN fused to a FTG on the N terminal and to an ETG on the C terminal. C. Same as B but with the expression cassette in reverse orientation.

doi:10.1371/journal.pone.0132502.g001

our GenoCAD plant grammar, we implemented a set of rules to direct the user through the design of plant expression vectors suitable for promoter studies (route *pro*; S2 Table).

The main requirement of promoter-reporter systems is that the sequence of the promoter of interest (native promoter) has to be fused to a reporter protein. To that end, the rule *npct* breaks down the expression cassette into a NPRO followed by a FTG and TER12. The rule *mpct* allows the user to clone the expression cassette in reverse orientation. Moreover, using the same rules mentioned above a single or double terminator can be added to the cassette, and epitope tags and/or linker domains can be included at both sides of the FTG. As an example, Fig 2 shows 2 different designs suitable for promoter analysis studies.

Since plant promoter information is still scarce, so far the compilation of native promoter sequences in the plant GenoCAD library is based on identification and characterization of several *Solanum tuberosum* promoters performed in our research group (e.g. StMCK6 promoter [21]). However, the idea is that each user expands the library with new entries of experimentally characterized promoters of interest [22].



**Fig 2. Example of two different designs for promoter analysis studies.** A. The expression cassette includes a FTG under the control of a NPRO and fused with an ETG by means of a LNK. B. The expression cassette has reverse orientation and double TER.

doi:10.1371/journal.pone.0132502.g002

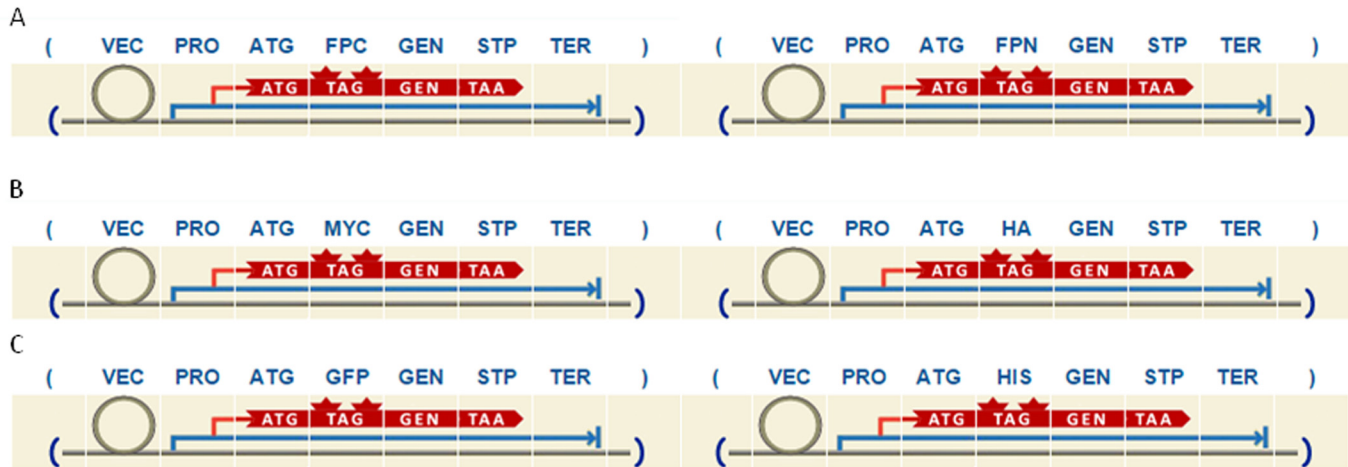
## Design expression plasmids for protein-protein interaction studies

PPI plays a crucial role in regulating biological systems, and provides valuable information about protein functions. In recent years, large-scale PPI studies have been applied to plants and allowed us to better understand the complex networks through which plant proteins exercise their functions (see reviews [23–25]). As an example, *Arabidopsis* interactome map [26] is the most complete interactome network published nowadays. Independently of how carefully large PPI experiments are performed, at least a representative subset of the data should be validated using an independent assay [27]. On the other hand, *in planta* functional validations of specific interactions are still necessary to understand the function of specific proteins of interest. Such experiments usually require the application of a combination of methods to a particular biological system. Over the last 10–15 years, many methods to characterize interactions in more realistic *in planta* settings have been developed (see review [24]). BiFC and targeted co-immunoprecipitation (CoIP) are the preferred choice for most of the researchers [28–30].

In BiFC approaches, two potential interacting proteins are fused to two fragments of a fluorescent protein, and identification of interactions is based on the reconstitution of the split fluorescent tag [31]. In the CoIP assay, the protein complex is precipitated with an immobilized antibody against one of the proteins studied, and the interacting partner is further confirmed by Western Blot. Usually antibodies against tagged fusion proteins are used [32]. According to the principle of these methodologies, the most important specificity of both studies is that 2 vectors have to be designed in parallel to allow proper experimental design. A split fluorescent protein is required in vectors for BiFC studies, and 2 different epitope tags in the case of the CoIP approach. To satisfy these requirements, the *ppi* route implemented in our grammar gives the users two options i.e. *bifc* and *coip* route (S2 Table). In the first case (Fig 3A), the rule defines that the design involves 2 plasmids, the first one containing the C-terminal of a half-fluorescent protein fused to the target protein (GFPC), and the second one containing the N-terminal of the same fluorescent protein (GFPN). Epitope tags and/or linker domains can also be added at both sides of the fluorescent protein (rules *tffc*, *fpct*, *lffc*, *fpcl*, *tfpn*, *fpnt*, *lfpn*, *fpnl*). On the other hand, the *coip* rule defines that the design involves 2 plasmids containing two different epitope tags fused to the potential interacting proteins. According to the most commonly used tags in plant molecular biology, two-pair combinations of epitope tags are included in the grammar i.e. c-myc (MYC)/Human influenza hemagglutinin (HA) and GFP/poly-histidine (HIS) tag. Specifically, rule *myha* defines that the first plasmid contains a cassette, which includes a gene fused to MYC epitope tag (GMY), and the second plasmid contains the gene fused to the HA epitope (GHA) (Fig 3B). In the same way, rule *gphi* forces the user to design the first plasmid with GFP fused to the gene (GGF), and the second plasmid with a gene fused to HIS tag (GHI) (Fig 3C). As in the case of localization studies, GEN category allow the user to add epitope tags and linkers at both sides of the protein under study, giving again high flexibility to the design.

## Discussion

We implemented a plant grammar in GenoCAD, which guides the user through the design of plant expression vectors for functional studies experiments. The PPI branch of the grammar includes rules that express the requirement that genetic parts located on different plasmids interact with one another. For instance, rules describing the use of split fluorescent proteins to visualize the interactions between two proteins require two plasmids to use sequences complementing each other. This type of systems-level analysis relying on trans-interactions between genetic parts had been proposed before [33] but this is the first time that this design constraint is captured in a grammar.



**Fig 3. Example of three different designs obtained following *ppi* route.** A) *bifc* route. B) *coip* route using MYC and HA as epitope tags. C) *coip* route using GFP and HIS as epitope tags.

doi:10.1371/journal.pone.0132502.g003

An important aspect of the grammar development effort was the development of a consistent set of icons suitable to properly represent the structure of the genetic constructs. We first attempted to use the icons of the Synthetic Biology Open Language (SBOL) in order to adhere to this emerging community standard [34]. SBOLv, the visualization component of the SBOL effort, currently includes 21 glyphs, 11 of which corresponding to functional elements as 10 icons are used to represent structural elements like restriction sites. Since our grammar includes 44 different categories, we could not find a satisfactory way to map our categories to the set of SBOLv icons. At some point, it seemed preferable to develop a set of custom icons. The icons have been designed to be consistent with one another so as to show the DNA, RNA, and proteins associated with different DNA sequences. The development of this grammar illustrates the conflicting requirements to develop a controlled vocabulary customized for a particular domain of application with the need to promote standards accepted by a large scientific community [35].

Even though it is possible to include in a grammar rules associated with specific assembly standards [22], rapid progress in DNA synthesis [36] makes it possible to assume users of this grammar would be able to obtain any sequence designed with the grammar by means of sequence-independent processes using the services of a DNA synthesis company or assembling these sequences in house [37].

The grammar we developed is based on our current experience in plant expression vectors. Rules were established according to the specifications required for specific plant functional analysis-oriented applications in order to design functional vectors that can be used for a wide range of plant species as long as they are transformable by *Agrobacterium tumefaciens*. We covered three types of applications i.e. protein localization, PPI and promoter analysis, which are indispensable and commonly used experiments among the plant biology community to discover and characterize gene functions.

The next step is now validation of constructs presented in the grammar since the complete final plasmids have not yet been experimentally verified for functionality. However, in our laboratory, all library parts have been tested in *Nicotiana benthamiana* using commercially available plasmids. As an example, for localization studies, the plant binary Gateway vector pH7YWG2 which allows the expression of translational fusions with YFP [38] was modified

and validated to express *Solanum tuberosum* mitogen-activated protein kinase (StWiPK) fused to mCherry.

The use of GenoCAD will certainly reduce time and cost of our experiments. For instance, the aforementioned simplest construct for localization studies (i.e. protein fused to a fluorescent tag) involves choosing between 3 different fluorescent tags at either N or C-terminal position bringing in 6 different plasmids. With the use of GenoCAD there is no need of a large amount of different commercial plasmids which meets all different experiment requirements. One single plasmid can be used to clone the different variant cassettes that better fulfill our needs with no plasmid modifications required and therefore fewer chances of errors during sequence manipulation.

The grammar is flexible enough to meet users' specific needs. It can generate constructs that may have questionable features, like a gene fused to several peptide linker sequences. We could have defined stricter constraints, such as imposing a linker sequence between protein domains. However, it has been recently reported that linker sequences are not critical factors for the success of BiFC experiments [29], therefore we decided to favor flexibility and let the user decide between different design options. If the need for additional features is identified by the user, the grammar can be revised and improved to encompass developments in the field.

In short, we presented a grammar implemented in GenoCAD to guide users through the design of plant expression vectors. The grammar is divided into different modules according to the aim/application of the final design. As an example, we developed rules for 3 of the main applications in plant functional studies: protein localization, promoter analysis, and PPI studies. We also curated a library of basic parts associated with the plant grammar, where each part is categorized into functional groups. The GenoCAD grammar editor makes it relatively easy to modify the grammar. A natural extension of our plant grammar would be to add rules for designing multigenic constructs, and also for other applications potentially interesting for plant biologists, e.g. silencing (VIGS, artificial microRNAs).

Taken altogether, this GenoCAD library will facilitate the design of large plant constructs that can be obtained using sequence-independent methods. It will contribute to transitioning methods from synthetic biology into mainstream plant science.

## Supporting Information

### S1 File. Plant Grammar.

(ZIP)

### S1 Table. Categories of genetic parts used in the plant grammar.

(XLS)

### S2 Table. The rules for plant grammar.

(XLS)

## Acknowledgments

This project is supported by NSF Awards 1241328, the Slovenian Research Agency Program P4-0165 and Project N4-0026.

## Author Contributions

Conceived and designed the experiments: AC MW JP KG. Performed the experiments: AC MW. Contributed reagents/materials/analysis tools: MW JP. Wrote the paper: AC MW JP KG.



## References

1. Silver P a, Way JC, Arnold FH. Synthetic biology: Engineering explored. *Nature*. 2014; 509: 166–167. doi: [10.1038/509166a](https://doi.org/10.1038/509166a) PMID: [24805338](https://pubmed.ncbi.nlm.nih.gov/24805338/)
2. Baltes NJ, Voytas DF. Enabling plant synthetic biology through genome engineering. *Trends Biotechnol*. Elsevier Ltd; 2014; 1–12. doi: [10.1016/j.tibtech.2014.11.008](https://doi.org/10.1016/j.tibtech.2014.11.008)
3. Patron NJ. DNA assembly for plant biology: techniques and tools. *Curr Opin Plant Biol*. Elsevier Ltd; 2014; 19C: 14–19. doi: [10.1016/j.pbi.2014.02.004](https://doi.org/10.1016/j.pbi.2014.02.004)
4. Czar MJ, Cai Y, Peccoud J. Writing DNA with GenoCAD. *Nucleic Acids Res*. 2009; 37: W40–7. doi: [10.1093/nar/gkp361](https://doi.org/10.1093/nar/gkp361) PMID: [19429897](https://pubmed.ncbi.nlm.nih.gov/19429897/)
5. Cai Y, Hartnett B, Gustafsson C, Peccoud J. A syntactic model to design and verify synthetic genetic constructs derived from standard biological parts. *Bioinformatics*. 2007; 23: 2760–7. doi: [10.1093/bioinformatics/btm446](https://doi.org/10.1093/bioinformatics/btm446) PMID: [17804435](https://pubmed.ncbi.nlm.nih.gov/17804435/)
6. Overend C, Yuan L, Peccoud J. The synthetic futures of vesicular stomatitis virus. *Trends Biotechnol*. Elsevier Ltd; 2012; 30: 497–8. doi: [10.1016/j.tibtech.2012.06.002](https://doi.org/10.1016/j.tibtech.2012.06.002) PMID: [22789133](https://pubmed.ncbi.nlm.nih.gov/22789133/)
7. Wilson ML, Okumoto S, Adam L, Peccoud J. Development of a domain-specific genetic language to design *Chlamydomonas reinhardtii* expression vectors. *Bioinformatics*. 2014; 30: 251–7. doi: [10.1093/bioinformatics/btt646](https://doi.org/10.1093/bioinformatics/btt646) PMID: [24215020](https://pubmed.ncbi.nlm.nih.gov/24215020/)
8. Purcell O, Peccoud J, Lu TK. Rule-Based Design of Synthetic Transcription Factors in Eukaryotes. *ACS Synth Biol*. 2014; 140103160806003. doi: [10.1021/sb400134k](https://doi.org/10.1021/sb400134k)
9. King SRF, McLellan H, Boevink PC, Armstrong MR, Bukharova T, Sukarta O, et al. Phytophthora infestans RXLR effector PexRD2 interacts with host MAPKKK  $\epsilon$  to suppress plant immune signaling. *Plant Cell*. 2014; 26: 1345–59. doi: [10.1105/tpc.113.120055](https://doi.org/10.1105/tpc.113.120055) PMID: [24632534](https://pubmed.ncbi.nlm.nih.gov/24632534/)
10. Tian Z, He Q, Wang H, Liu Y, Zhang Y, Shao F, et al. The Potato ERF Transcription Factor STERF3 Negatively Regulates Resistance to Phytophthora infestans and Salt Tolerance in Potato. *Plant Cell Physiol*. 2015; 56: 992–1005. doi: [10.1093/pcp/pcv025](https://doi.org/10.1093/pcp/pcv025) PMID: [25681825](https://pubmed.ncbi.nlm.nih.gov/25681825/)
11. Hu Y, Chen L, Wang H, Zhang L, Wang F, Yu D. Arabidopsis transcription factor WRKY8 functions antagonistically with its interacting partner VQ9 to modulate salinity stress tolerance. *Plant J*. 2013; doi: [10.1111/tpj.12159](https://doi.org/10.1111/tpj.12159)
12. Chen X, Zaro JL, Shen W-C. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev*. Elsevier B.V.; 2013; 65: 1357–69. doi: [10.1016/j.addr.2012.09.039](https://doi.org/10.1016/j.addr.2012.09.039) PMID: [23026637](https://pubmed.ncbi.nlm.nih.gov/23026637/)
13. The Potato Genome Sequencing Consortium. Genome sequence and analysis of the tuber crop potato. *Nature*. 2011; 475: 189–95. doi: [10.1038/nature10158](https://doi.org/10.1038/nature10158) PMID: [21743474](https://pubmed.ncbi.nlm.nih.gov/21743474/)
14. Voss U, Larrieu A, Wells DM. From jellyfish to biosensors: the use of fluorescent proteins in plants. *Int J Dev Biol*. 2013; 57: 525–33. doi: [10.1387/ijdb.130208dw](https://doi.org/10.1387/ijdb.130208dw) PMID: [24166435](https://pubmed.ncbi.nlm.nih.gov/24166435/)
15. Kerppola TK. Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nat Protoc*. 2006; 1: 1278–86. doi: [10.1038/nprot.2006.201](https://doi.org/10.1038/nprot.2006.201) PMID: [17406412](https://pubmed.ncbi.nlm.nih.gov/17406412/)
16. Walter M, Chaban C, Schütze K, Batistic O, Weckermann K, Näge C, et al. Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *Plant J*. 2004; 40: 428–438. doi: [10.1111/j.1365-313X.2004.02219.x](https://doi.org/10.1111/j.1365-313X.2004.02219.x) PMID: [15469500](https://pubmed.ncbi.nlm.nih.gov/15469500/)
17. Kerppola TK. Visualization of molecular interactions using bimolecular fluorescence complementation analysis: characteristics of protein fragment complementation. *Chem Soc Rev*. 2009; 38: 2876–86. doi: [10.1039/b909638h](https://doi.org/10.1039/b909638h) PMID: [19771334](https://pubmed.ncbi.nlm.nih.gov/19771334/)
18. Lee L-Y, Gelvin SB. T-DNA binary vectors and systems. *Plant Physiol*. 2008; 146: 325–332. doi: [10.1104/pp.107.113001](https://doi.org/10.1104/pp.107.113001) PMID: [18250230](https://pubmed.ncbi.nlm.nih.gov/18250230/)
19. Runions J, Hawes C, Kurup S. Fluorescent Protein Fusions for Protein Localization in Plants. In: Van der Giezen M, editor. *Protein Targeting Protocols; Methods in Molecular Biology Volume 390*. 2nd Edition. 2007. pp. 239–255.
20. Pan Y, Ma X, Liang H, Zhao Q, Zhu D, Yu J. Spatial and temporal activity of the foxtail millet (*Setaria italica*) seed-specific promoter pF128. *Planta*. 2014; doi: [10.1007/s00425-014-2164-5](https://doi.org/10.1007/s00425-014-2164-5)
21. Lazar A, Coll A, Dobnik D, Baebler S, Bedina-Zavec A, Zel J, et al. Involvement of Potato (*Solanum tuberosum* L.) MKK6 in Response to Potato virus Y. *PLoS One*. 2014; 9: e104553. doi: [10.1371/journal.pone.0104553](https://doi.org/10.1371/journal.pone.0104553) PMID: [25111695](https://pubmed.ncbi.nlm.nih.gov/25111695/)
22. Cai Y, Wilson ML, Peccoud J. GenoCAD for iGEM: a grammatical approach to the design of standard-compliant constructs. *Nucleic Acids Res*. 2010; 38: 2637–44. doi: [10.1093/nar/gkq086](https://doi.org/10.1093/nar/gkq086) PMID: [20167639](https://pubmed.ncbi.nlm.nih.gov/20167639/)
23. Fukao Y. Protein-protein interactions in plants. *Plant Cell Physiol*. 2012; 53: 617–25. doi: [10.1093/pcp/pcs026](https://doi.org/10.1093/pcp/pcs026) PMID: [22383626](https://pubmed.ncbi.nlm.nih.gov/22383626/)

24. Braun P, Aubourg S, Van Leene J, De Jaeger G, Lurin C. Plant Protein Interactomes. *Annu Rev Plant Biol.* 2013; 1–27. doi: [10.1146/annurev-arplant-050312-120140](https://doi.org/10.1146/annurev-arplant-050312-120140)
25. Zhang Y, Gao P, Yuan JS. Plant protein-protein interaction network and interactome. *Curr Genomics.* 2010; 11: 40–6. doi: [10.2174/138920210790218016](https://doi.org/10.2174/138920210790218016) PMID: [20808522](https://pubmed.ncbi.nlm.nih.gov/20808522/)
26. Arabidopsis Interactome Mapping Consortium. Evidence for network evolution in an Arabidopsis interactome map. *Science.* 2011; 333: 601–7. doi: [10.1126/science.1203877](https://doi.org/10.1126/science.1203877) PMID: [21798944](https://pubmed.ncbi.nlm.nih.gov/21798944/)
27. Dreze M, Monachello D, Lurin C, Cusick ME, Hill DE, Vidal M, et al. High-quality binary interactome mapping. *Methods Enzymol.* 2010; 470: 281–315. doi: [10.1016/S0076-6879\(10\)70012-4](https://doi.org/10.1016/S0076-6879(10)70012-4) PMID: [20946815](https://pubmed.ncbi.nlm.nih.gov/20946815/)
28. Kodama Y, Hu CD. Bimolecular fluorescence complementation (BiFC): A 5-year update and future perspectives. *Biotechniques.* 2012; 53: 285–298. doi: [10.2144/000113943](https://doi.org/10.2144/000113943) PMID: [23148879](https://pubmed.ncbi.nlm.nih.gov/23148879/)
29. Horstman A, Tonaco IAN, Boutillier K, Immink RGH. A Cautionary Note on the Use of Split-YFP/BiFC in Plant Protein-Protein Interaction Studies. *Int J Mol Sci.* 2014; 15: 9628–43. doi: [10.3390/ijms15069628](https://doi.org/10.3390/ijms15069628) PMID: [24886811](https://pubmed.ncbi.nlm.nih.gov/24886811/)
30. Speth C, Toledo-Filho LAA, Laubinger S. Immunoprecipitation-Based Analysis of Protein–Protein Interactions. In: Staiger D, editor. *Plant Circadian Networks; Methods in Molecular Biology.* Springer New York; 2014. pp. 175–185.
31. Kerppola TK. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu Rev Biophys.* Annual Reviews; 2008; 37: 465–87. doi: [10.1146/annurev.biophys.37.032807.125842](https://doi.org/10.1146/annurev.biophys.37.032807.125842)
32. Ransone LJ. Oncogene Techniques: Detection of protein-protein interactions by coimmunoprecipitation and dimerization [Internet]. *Methods in Enzymology.* Elsevier; 1995. pp. 491–497. doi: [10.1016/0076-6879\(95\)54034-2](https://doi.org/10.1016/0076-6879(95)54034-2) PMID: [8531709](https://pubmed.ncbi.nlm.nih.gov/8531709/)
33. Ball D a, Lux MW, Graef RR, Peterson MW, Valenti JD, Dileo J, et al. Co-design in synthetic biology: a system-level analysis of the development of an environmental sensing device. *Pacific Symposium on Biocomputing.* 2010. pp. 385–396. 9789814295291\_0041 [pii]
34. Galdzicki M, Clancy KP, Oberortner E, Pocock M, Quinn JY, Rodriguez C a, et al. The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nat Biotechnol.* 2014; 32: 545–550. doi: [10.1038/nbt.2891](https://doi.org/10.1038/nbt.2891) PMID: [24911500](https://pubmed.ncbi.nlm.nih.gov/24911500/)
35. Adames NR, Wilson ML, Fang G, Lux MW, Glick BS, Peccoud J. GenoLIB: a database of biological parts derived from a library of common plasmid features. *Nucleic Acids Res.* 2015; 1–10. doi: [10.1093/nar/gkv272](https://doi.org/10.1093/nar/gkv272)
36. Czar MJ, Anderson JC, Bader JS, Peccoud J. Gene synthesis demystified. *Trends Biotechnol.* 2009; 27: 63–72. doi: [10.1016/j.tibtech.2008.10.007](https://doi.org/10.1016/j.tibtech.2008.10.007) PMID: [19111926](https://pubmed.ncbi.nlm.nih.gov/19111926/)
37. Peccoud J, editor. *Gene Synthesis: Methods and Protocols.* Methods in Molecular Biology. Springer Science + Business Media; 2012. p. 297. doi: [10.1007/978-1-62703-239-1\\_1](https://doi.org/10.1007/978-1-62703-239-1_1)
38. Karimi M, De Meyer B, Hilson P. Modular cloning in plant cells. *Trends Plant Sci.* 2005; 10: 103–105. doi: [10.1016/j.tplants.2005.01.008](https://doi.org/10.1016/j.tplants.2005.01.008) PMID: [15749466](https://pubmed.ncbi.nlm.nih.gov/15749466/)