Journal of Plant Genetics and Breeding

White, J Plant Genet Breed 2022, 6:6

Research Article Open Access

GC Content of Plant Genes is Linked to Past Gene Duplications

John White*

Instituto de Investigaciones Biomédicas, Universidad de las Américas, Quito, Ecuador

Abstract

From species to species, and occasionally even between different genes in the same genome, the frequency of G and C nucleotides varies in genomes. In contrast to dicots, monocot grasses have a bimodal distribution of genic GC content. Syntenic genes had significantly higher GC content than non-syntenic genes at their 5'-end in the third position among codons for all 9 species, according to our classification of plant genes from 5 dicots and 4 monocot grasses to related species. Gene duplication is associated with lower GC content because old scattered gene duplications lack synteny to distantly related genomes. Biased GC content can be explained by two mutation types: methylation C to T mutation and gene conversion from A to G. Gene duplication to additional loci is likely to result in non-reciprocal exchanges between homologous alleles, which make it impossible to distinguish when the alleles are identical or heterozygous for presence-absence variation. Gene duplication can result in siRNA synthesis, which can lead to targeted methylation and increase mC-T mutations. Plant genes that have recently been duplicated are more frequently methylated and less likely to go through gene conversion, which together provide a mutational environment that favors AT nucleotides [1]. A subgroup of the syntenic genes in grasses with high GC content has undergone few duplications or has had its duplicate copies eliminated by selection. We put forth the "biased gene duplication / biased mutation" (BDBM) theory as a potential explanation for the genesis and evolution of the duplication-genic GC bias relationship. Empirical evidence for the BDBM model is provided by joint analyses of the genes from 9 angiosperm species, with the genes' duplication status, GC content, methylation levels, and functional classes categorized.

Keywords: Angiosperm; Pectin polysaccharides; GC content; Gene duplication; Cell adhesion

Introduction

Pectin polysaccharides, which make up a large fraction of the immediate contact between cells, are necessary for cell adhesion in plants. The main cell wall is produced through a combination of cellulose secretion at the plasma membrane and secretion through the endomembrane system of pectin, hemicellulose, and other polysaccharides, which is where the cell wall is initially deposited during cell division [2]. The rigidity and elasticity of the cell wall can be significantly affected by enzymatic activity that further modifies the polymers. For instance, pectin methyl esterase (PME) mediates the de-esterification of pectin, which is inhibited by PME inhibitors (PMEI), which controls the charge and calcium dependent crosslinking and is related to changes in wall extensibility and adhesion. The degree of pectin esterification can affect the activity of polygalacturonases and pectate lyases, which results in a complex interplay of enzyme expression profiles and substrate/enzyme pairing across various tissue types and plant species [3].

Cell adhesion can be lost as a result of mutations that either decrease or modify the pectin content or alter their modification. A 50% reduction in pectin and a considerable loss of cell adhesion are caused by mutations in QUASIMODO 1 and 2 (QUA1, 2), which encode a Golgi-localized glycosyl and methyl transferase, respectively. Hypocotyl expansion is the easiest place to detect these effects. However, mutations in FRIABLE 1 (FRB1), which also results in a similar adhesion defect, alter the amount of oligosaccharides in the Golgi that contain galactose and arabinose, as well as the methyl esterification of pectin and the microstructure of xyloglucan, but they have no effect on the total amount of pectin. However, a signalling cascade mediated by the putative O-fucosytransferase ESMERALDA1 (ESMD1), which is thought to use an EGF domain as a substrate for the addition of a single fucose to a serine or threonine at the consensus C2XXXXS/TC3 (where X is any amino acid and numbers indicate one of six cysteine's,), may also regulate cell adhesion EGF domains are distinguished by a series of six repeated and evenly spaced cysteine's, and their fucosylation can change the receptor activation in metazoans [4]. Qua2-1 and frb1 can be suppressed by mutations in ESMD1, but pectin levels are not restored. The characteristics of this hypothetical signalling pathway are unknown, but plant receptor kinases with an EGF domain may hold information about it.

Arabidopsis has at least two receptor protein families that have EGF domains with potential O-fucosylation sites. Each member of the six-member family of G-type lectin S-receptor-like serine/threonine-protein kinases has an EGF domain with a consensus O-fucosylation sequence. To date, there have been no findings that connect these kinds of receptors to pectin or the cell wall [5]. However, the pectin-binding Wall associated kinases (WAKs) also have a potential ESMD1 substrate, and because there is strong evidence linking WAKs to pectin signaling, they are the top contenders for ESMD1-mediated O-fucosylation. A cytoplasmic serine/threonine kinase and many extracellular EGF domains are features of the protein kinases known as WAKs. In addition to binding to native cell wall pectin, which is necessary for cell growth, WAKs can attach to pectin fragments or oligogalacturonides (OGs), which can be produced as a result of injury, pathogen infection, dark/light exposure, or stress. Shorter OGs can also participate in photomorphogenesis. Although it is unknown how

***Corresponding author:** John white, Instituto de Investigaciones Biomédicas, Universidad de las Américas, Quito, Ecuador, E-mail: jwhite@uga.edu

Received: 01-Nov-2022, Manuscript No: jpgb-22-82727, **Editor Assigned:** 04-Nov-2022, pre QC No: jpgb-22-82727 (PQ), **Reviewed:** 18-Nov-2022, QC No: jpgb-22-82727, **Revised:** 25-Nov-2022, Manuscript No: jpgb-22-82727 (R), **Published:** 30-Nov-2022, DOI: 10.4172/jpgb.1000131

Citation: White J (2022) GC Content of Plant Genes is Linked to Past Gene Duplications. J Plant Genet Breed 6: 131.

Copyright: © 2022 White J. 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.

WAKs differentiate these OGs from native pectin, it has been proposed that long polymers and fragmented pectin compete for WAK activation to activate alternative pathways [6].

Material and Methods

Plant Species, Growing Environment, And Strain Information

In this investigation, transformation was carried out on the P. sojae-resistant soybean cultivar "Williams 82," which carries the Rps1k resistance gene. In a growth environment with a 14-h photoperiod, a relative humidity of 70 10%, day/night temperatures of 22°C/18°C, and a light intensity of 350 mol m-2s-1, seeds obtained from the T0 generation were planted in pots filled with sterile vermiculite [7].

The dominating race in Jilin Province, P. sojae race 1, PSR01, was graciously donated by Professor Shuzhen Zhang and her team. In Heilongjiang, it was isolated from diseased soybean plants.

Designing siRNA and building CRISPR/Cas9 expression vectors

The Phytozome database was used to download the sequence for the soybean endogenous gene GmTCP19L (Glyma.05G050400.1). The CRISPR-P web tool was used to find potential siRNA target locations within the GmTCP19L gene. Using Primer Premier 5.0, the primer binding sites for the amplification of particular siRNA target sites were created. CDSearch was used to forecast the functional domain [8]. The sequence of Cas9 was assembled downstream of the CaMV 35S promoter along with the siRNA driven by the Glycine max U6 promoter (GmU6) within its T-DNA region to create the GmTCP19L-CRISPR/Cas9 vector, carrying both GmTCP19L targeted siRNA and Cas9 cassettes. The bar gene driven by a CaMV 35S promoter was used as a screening marker.

Sequencing analysis for mutation detection

In the T0, T1 and T2 generations, genomic DNA was isolated from each plant's leaves using the modified cetyltrimethylammonium ammonium bromide (CTAB) procedure. Next, the 621 bp GmTCP19L target area was amplified by PCR using Phanta Super Fidelity DNA Polymerase. A 1% agarose gel electrophoresis was used to find the PCR products and the GmTCP19L-F and GmTCP19L-R primers were used to sequence the results. Sequence peaks can distinguish between three different kinds of gene editing. While the wild-type (WT) and homozygous mutations had distinct peaks at the target location, the heterozygous mutations displayed overlapping peaks [9]. Sequence alignment with the WT sequence allowed for the detection of the homozygous mutations, and CRISPR/Cas9-induced frameshift mutations can result from small base insertions or deletions. We simultaneously screened T1 and T2 progenies for the tcp19l mutants devoid of the BAR and Cas9 sequences of the CRISPR/Cas9 vector. Following the manufacturer's instructions, the PAT/Bar test strip was used to identify the BAR protein, and the primers Cas9-F/R were used to amplify the fragment (349 bp) of the Cas9 gene [10].

Examining off-target mutagenesis

We used the online website tool CRISPR-P to analyse the potential off-target sites in order to determine whether either of the target sites might exhibit off-target activity. The sequences for the two GmTCP19L-SP1 and GmTCP19L-SP2 regions with the highest potential for off-target activity were downloaded from the Ensembl Plants database. Each off-target site's primers were created to amplify 300–500 bp regions. Through the use of the PCR technique, the regions spanning the target sites were amplified. Sequencing analysis was then used to determine the various kinds of potential off-target site editing. Detection of enzyme activities.

To test whether GmTCP19L could affect superoxide dismutase (SOD) activity and peroxidase (POD) activity, the activities of SOD and POD were measured in tcp19l mutants one gram of fresh roots were harvested at 3 days after inoculation with zoospores of P. sojae. The SOD and POD activities were measured as described previously by Li et al. (2015) [11]. Non-transformed seedlings were used as controls.

Discussion

Biased mutation and biased gene duplication

The relationship between GC3-50 and the synteny status of plant genes, particularly in grasses, points to a possible mechanism or pathways connecting these two traits. Individual genes can undergo non-tandem duplication, in which case the new copy is neither syntenic nor redundant, and the original (syntenic) copy can be lost by mutation with little to no effect on fitness. The likelihood that members of a gene family are not syntenic increases with repeated gene duplications and loss in either lineage. A subset of the gene space that is less duplication prone, or situations in which the duplicate copies were chosen against and swiftly lost, presumably due to gene dosage balance, is represented by genes that are still syntenic with distantly related genomes [13]**.** Therefore, compared to genes that have stayed single copy and syntenic, plant genes that have undergone more non-tandem gene duplications have encountered a distinct mutational environment.

Since most alterations to the third base of a codon do not affect the encoded protein, the bimodal distribution of GC content in grasses is unlikely to be the result of selection. While there is a substantial association between the GC content of grass genes and gene expression levels [8] and/or tissue specificity, the effect seems to be quite small. While selection is probably only going to have a little impact, GC content is the main determinant of codon bias in a variety of species [14].

The majority of observed SNPs in plant genomes are often explained by C–T transitions. For instance, 52.1% of SNPs in Arabidopsis and 70% of SNPs in rice are C/T polymorphisms. In the absence of selection, the C/T mutation rates are the main determinant of C/G frequency since, in rice and Arabidopsis, respectively, 15% and 25.2% of other SNPs contain A/T or C/G polymorphisms. Since gene conversion and higher mutation brought on by methylation cytosine both increase C-T polymorphisms, one or both of these mechanisms—or both—could be responsible for the observed regional difference in GC content [15].

Conclusion

In the field of animal research, the regulation of RHO family proteins by PTMs has long been known. But the study of these PTM regulation mechanisms in ROP signalling in plants is still in its infancy. Here, we report information on barley ROP's in vivo ubiquitination. The discovered ubiquitination site has previously been characterized for mammalian RACB-homologs and is conserved in all ROPs from barley, rice, and Arabidopsis. This shows that ubiquitination affects protein abundance and that the lysine residue corresponding to RACB K167 is a common target across kingdoms.

Acknowledgement

None

Conflict of Interest

None

References

1. Luo D, Carpenter R, Vincent C, Copsey L, Coen E (1996) Origin of floral asymmetry in Antirrhinum. Nature 383: 794-799.
2. Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. Nature 386: 485-488.
3. Cubas P, Lauter N, Doebley J, Coen E (1999) The TCP domain: a motif found in proteins regulating plant growth and development. Plant J 18: 215-222.
4. Kosugi S, Ohashi Y (2002) DNA binding and dimerization specificity and potential targets for the TCP protein family. Plant J 30: 337-348.
5. Navaud O, Dabos P, Carnus E, Tremousaygue D, Hervé C (2007) TCP transcription factors predate the emergence of land plants. J Mol Evol 65: 23-33.
6. Hervé C, Dabos P, Bardet C, Jauneau A, Auriac MC, et al. (2009) In vivo interference with AtTCP20 function induces severe plant growth alterations and deregulates the expression of many genes important for development. Plant Physiol 149: 1462-1477.
7. Li DB, Zhang HY, Mou MH, Chen YL, Xiang SY, et al. (2019) Arabidopsis class II TCP transcription factors integrate with the FT–FD module to control flowering. Plant Physiol 181: 97–111.
8. Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, et al. (2003) Control of leaf morphogenesis by micro RNAs. Nature 425: 257-263.
9. Balsemão-Pires E, Andrade L, Sachetto-Martins G (2013) Functional study of TCP23 in Arabidopsis thaliana during plant development. Plant Physiol Biochem 67: 120-125.
10. Danisman S, van Dijk A, Bimbo A, van der Wal F, Hennig L, et al. (2013) Analysis of functional redundancies within the Arabidopsis TCP transcription factor family. J Exp Bot 64: 5673-5685.
11. Sugio A, MacLean AM, Hogenhout SA (2014) The small phytoplasma virulence effector SAP11 contains distinct domains required for nuclear targeting and CIN-TCP binding and destabilization. New Phytol 202: 838-848.
12. Colas M, Cubas P (2016) TCP factors: New kids on the signaling block. Curr Opin Plant Biol 33: 33-41.
13. Tomotsugu K, Fumihiko S, Masaru OT (2017) Roles of miR319 and TCP transcription factors in leaf development. Plant Physiol 175: 874-885.
14. Lei N, Yu X, Li SX, Zeng CY, Zou LP, et al. (2017) Phylogeny and expression pattern analysis of TCP transcription factors in cassava seedlings exposed to cold and/or drought stress. Sci Rep 7: 10016.
15. Li W, Li DD, Han LH, Tao M, Hu QQ, et al. (2017) Genome-wide identification and characterization of TCP transcription factor genes in upland cotton (Gossypium hirsutum). Sci Rep 7: 10118.