IUNS 21st ICN
International Congress of Nutrition
"From Sciences to Nutrition Security"
Buenos Aires, Argentina, 15-20 October 2017 - Sheraton Buenos Aires Hotel & Convention Center
www.iuns-icn2017.com info@iuns-icn2017.com

SAN Sociedad Argentina de Nutrición
IUNS INTERNATIONAL UNION OF NUTRITIONAL SCIENCES

Reference
144/2106

Insights from the Sequencing and Annotation of the *Stevia rebaudiana* Genome and their Application in Agronomy and Health

Stephen E. Schauer PhD[1], Runchun Jing PhD[2], Ong Seong Siang PhD[2], Wong Yeen Yee PhD[2], Yucheng Bu[2], Jianning Chen[2], Jifeng Tang PhD[1], Raymond van Daelen PhD[1], Tengfang Huang PhD[1], Indra Prakash PhD[3], Priscilla Samuel PhD[4], Antoine Janssen MSc[1], Marcel van Verk PhD[1], Fayaz Khazi PhD[1], Walter Nelson MBA[1], Alec Hayes PhD[3], Avetik Markosyan MD, PhD[2]
[1]KeyGene, [2]PureCircle Ltd, [3]The Coca-Cola Company, [4]Global Stevia Institute

Background and Objectives:

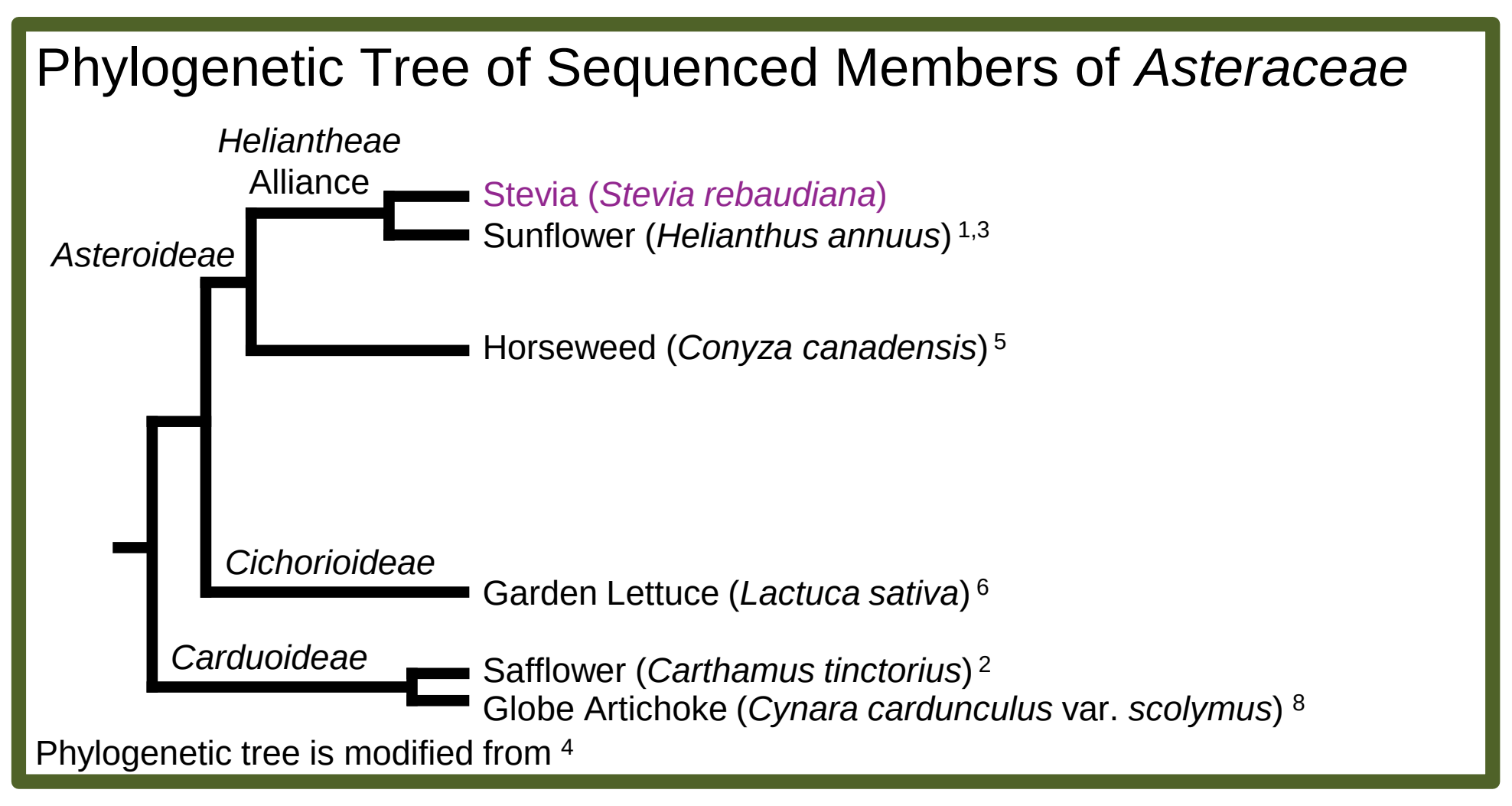
Obesity is a global health issue affecting millions and reducing sugar intake with nonnutritive sweeteners is one potential path to reducing dietary energy consumption that can have a positive health impact. The leaves of the *Stevia rebaudiana* (Bertoni) plant produce steviol glycosides: including stevioside; rebaudiosides A, B, C, D, E, F, and M; dulcoside A; rubusoside; and steviolbioside. Steviol glycosides are 200-350 times sweeter than sugar but contribute no calories, therefore making Stevia a leaf crop of significant economic value. The food industry has a very positive outlook on the opportunities for Stevia ingredients in zero or calorie-reduced products. Consequently, the Stevia ingredient market is expected to cross $1 billion USD by the year 2021[10]. High yields; desired quality traits; resistance to pests and diseases; and dependable economic returns are key characteristics of any domesticated commercial crop. Such basic characteristics are generally selected over many generations through intensive plant breeding efforts to create improved varieties. Although systematic cultivation of Stevia started in the 1970s in China, South America and Japan, only recently crop improvement efforts have been focused on making Stevia more scalable and an economically sustainable crop.

Methods:

Despite limited understanding of Stevia's genetics, biology or plant physiology, recent demand for steviol glycosides requires rapid improvement of Stevia cultivars currently used in commercial cultivation. To rapidly facilitate this, we have sequenced, generated genome assemblies, and fully annotated the genomes of three commercial Stevia varieties with improved levels of the better tasting, more sugar-like minor rebaudiosides.

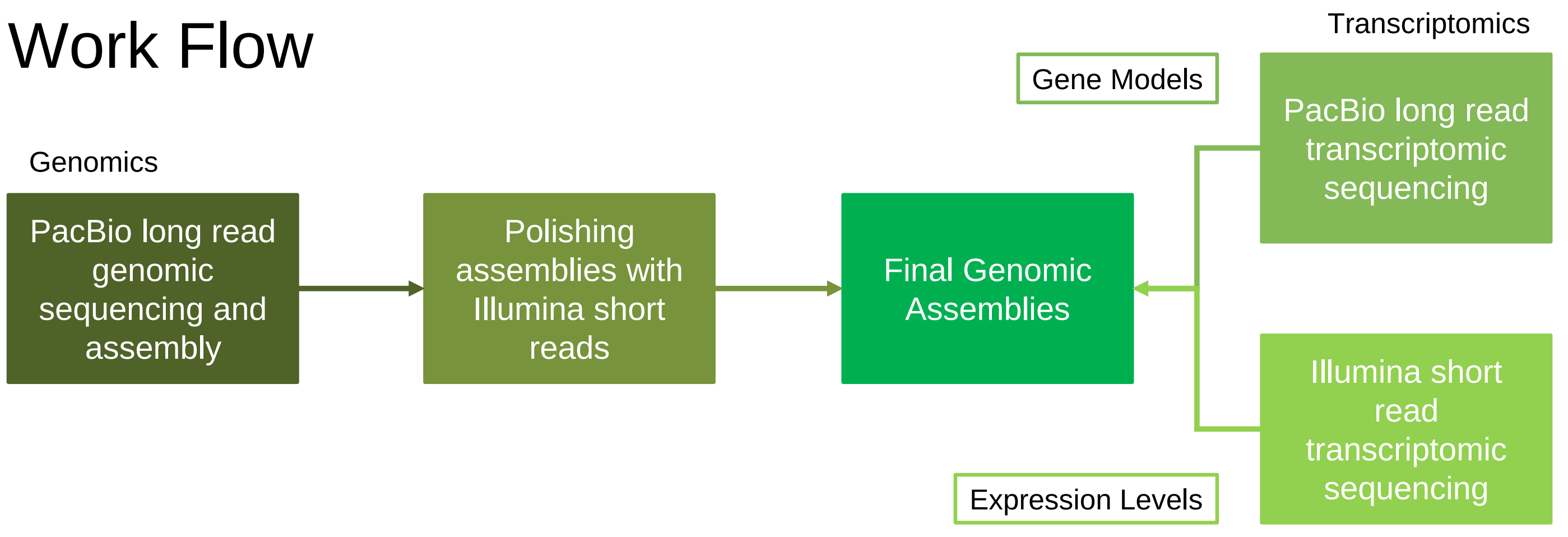

Stevia rebaudiana

Phylogenetic Tree of Sequenced Members of *Asteraceae*



Phylogenetic tree is modified from [4]

Work Flow



Results:

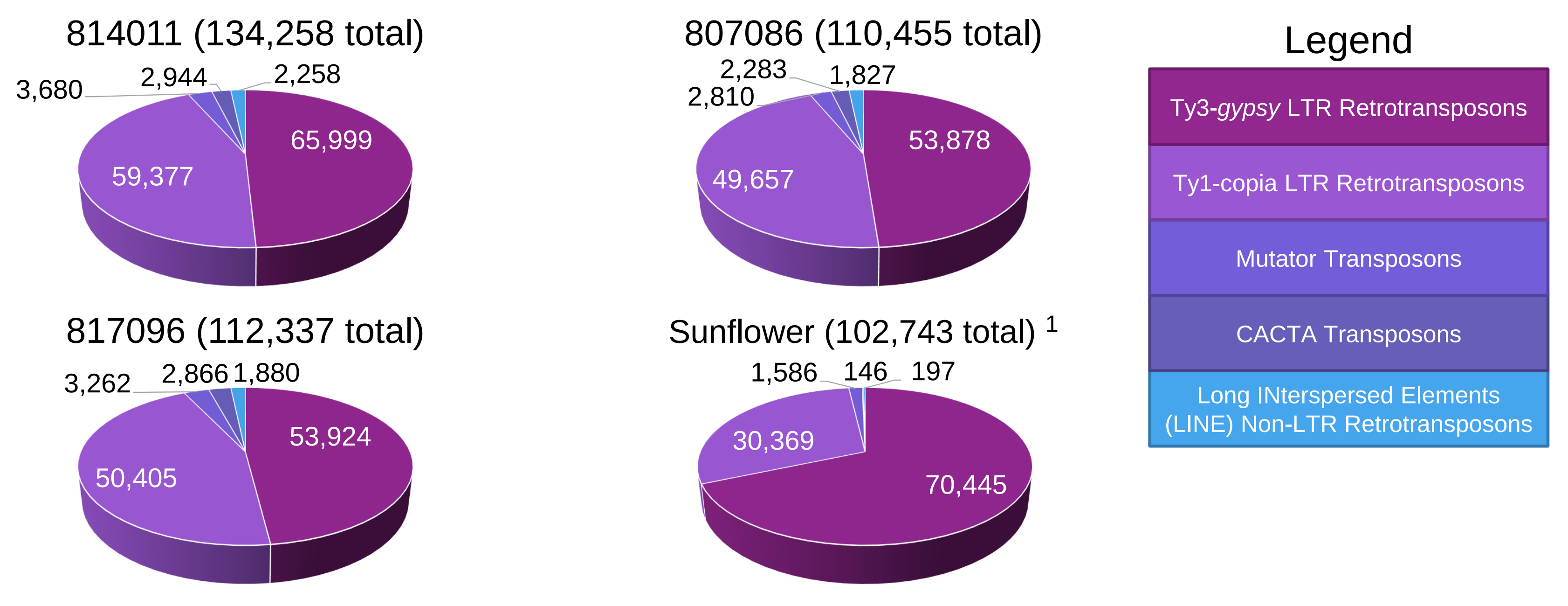
These data have been integrated into a comprehensive bioinformatics platform for visualization and analytics of all available genomic, transcriptomic and metabolomic Stevia datasets. This interface will enable specialists from multiple disciplines, such as chemists, biochemists and geneticists, to mine this platform to understand and improve existing steviol glycoside biosynthesis pathways through traditional breeding or discover new pathways or compounds for similar non-GMO improvement.

Assembly Metrics for Stevia And Other Sequenced Members of the *Asteraceae*

Metrics	Stevia			Sunflower		Horseweed	Lettuce	Safflower	Artichoke
Variety References	814011	807086	817096	HA412-HO [3]	XRQ [1]	TN-R [5]	Salinas [6]	AC Sunset [2]	2C [8]
Predicted Haploid Genome Size in Gigabase-pairs (1Gb = 10^9)	1.15 ± 0.03	1.17 ± 0.01	1.20 ± 0.01	3.6		0.34	2.5	1.35	1.08
Long Read (PacBio) Coverage of Haploid Genome	75X	71X	81X	—	102X	4.5X	—	—	—
Short Read Coverage of Haploid Genome / Sequencing Platform Used	151X / Illumina	221X / Illumina	125X / Illumina	80X / Solexa 15X / 454	—	230X / Illumina	72.5X / Illumina	21X / Illumina	133X / Illumina
Total Gbs in assembly	1.7	1.4	1.4	~2.0	2.9	0.31	2.21	0.87	0.66
Total Number of Contigs (Contiguous DNA fragments)	9,814	7,071	14,405	~800,000	12,318	20,075	21,116	3,254,412	79,681
Median Contig Length (N_{50}) in Kilobase-pairs (1Kb = 10^3)	332	441	184	—	524	21	36	0.4	18
Number of Predicted Gene Models	91,623	74,534	74,785	77,855	58,035	44,592	31,348	—	26,889
BUSCO[9] genes detected; a measure of completeness	89%	89%	88%	—	92%	—	—	—	—

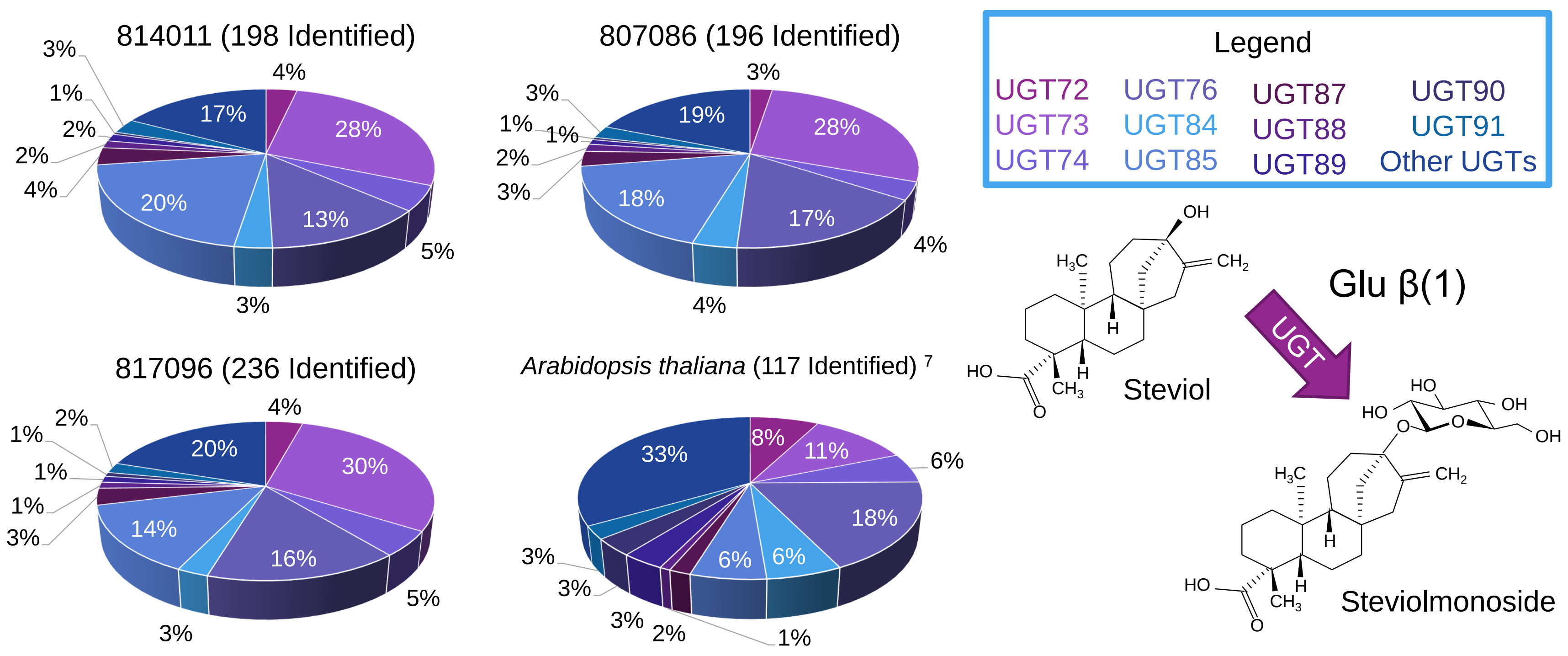
The assembled genomes for the three Stevia varieties are comparable to the other published genomes from *Asteraceae*

High Frequency of Retrotransposons Present in *Asteroideae*



Repetitive elements, such as transposons and retrotransposons, are a common feature of all genomes. Both the closely related Stevia and Sunflower appear to have a large percentage of their repetitive elements as retrotransposons, mainly of the *gypsy* and *copia* families. Large number of repetitive elements such as found in the *Asteroideae* increase the complexity of genome assembly.

Expansion of UDP-Glycosyltransferase (UGT) Families in Stevia



UGTs mediate the transfer of a glycosyl group to acceptor molecules, an enzymatic process that is critical for the production of steviol glycosides, amongst many additional secondary metabolites. In Stevia, this class of enzymes appears to be expanded as compared to other plants (such as the well known model plant *Arabidopsis*), which may affect the production of steviol glycosides.

Conclusions:

Here the assembly of the genomes of three Stevia varieties is presented, which are all comparable to other published high quality genomes of the *Asteraceae*. There appears to be expansion of the UGT family of enzymes, which play a critical role in the production of steviol glycosides; this expansion might help explain the diverse set of steviol glycosides found in Stevia. In addition to enabling the improvement of traditional breeding for agronomic and sustainability benefits, this new knowledge of the Stevia genome can facilitate improvements in the abundance of the steviol glycosides that are more sugar-like in taste, thus enabling the development of high-purity Stevia leaf extracts for deeper reductions in sugar and calories in food and beverage products. Stevia leaf extracts are natural-origin, plant-based, zero-calorie sustainable sweeteners, thus appealing to those looking for more natural plant-based ingredients in their diet. Food and beverage products with reduced sugar and calories can help consumers meet personal and public health priorities for improved diets and health.

References:

1. Badouin *et al.*, (2017) The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature*. **546:** 148-152.
2. Bowers *et al.*, (2016) Genetic Mapping of Millions of SNPs in Safflower (*Carthamus tinctorius* L.) via Whole-Genome Resequencing. *G3 (Bethesda)*. **6:** 2203-2211.
3. Kane *et al.*, (2011) Progress towards a reference genome for sunflower. *Botany*. **89:** 429-437.
4. Panero and Funk (2008) The value of sampling anomalous taxa in phylogenetic studies: major clades of the Asteraceae revealed. *Mol Phylogenet Evol.* **47:** 757-782.
5. Peng *et al.*, (2014) De novo genome assembly of the economically important weed horseweed using integrated data from multiple sequencing platforms. *Plant Physiol.* **166:** 1241-1254.
6. Reyes-Chin-Wo *et al.*, (2017) Genome assembly with in vitro proximity ligation data and whole-genome triplication in lettuce. *Nat Commun.* **8:**14953. doi: 10.1038/ncomms14953.
7. Ross *et al.*, (2001) Higher plant glycosyltransferases. *Genome Biol.* **2:** REVIEWS3004.
8. Scaglione *et al.*, (2016) The genome sequence of the outbreeding globe artichoke constructed *de novo* incorporating a phase-aware low-pass sequencing strategy of F1 progeny. *Sci Rep.* **6:** 19427.
9. Simão *et al.*, (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31:** 3210-3212.
10. International Stevia Council (2017) Industry data report, September 2017.

Keywords:

Stevia, Genome, Breeding, Agronomy, Steviol Glycosides

Conflict of interest:

Stephen E. Schauer, Jifeng Tang, Raymond van Daelen, Tengfang Huang, Antoine Janssen, Marcel van Verk, Fayaz Khazi, and Walter Nelson work or worked at KeyGene. Runchun Jing, Ong Seong Siang, Wong Yeen Yee, Yucheng Bu, Jianning Chen, and Avetik Markosyan work at PureCircle Ltd. Indra Prakash and Alec Hayes work at The Coca-Cola Company. Priscilla Samuel works at the Global Stevia Institute which is funded by PureCircle Ltd. PureCircle Ltd and The Coca-Cola Company co-funded the genome research.

PureCircle Everything stevia
KeyGene
Coca-Cola
Global Stevia Institute