



OPEN

# Unprecedented enhancement of recombinant protein production in sugarcane culms using a combinatorial promoter stacking system

Mona B. Damaj<sup>1</sup>✉, John L. Jifon<sup>1,2</sup>, Susan L. Woodard<sup>3</sup>, Carol Vargas-Bautista<sup>1,7</sup>, Georgia O. F. Barros<sup>4</sup>, Joe Molina<sup>1</sup>, Steven G. White<sup>4</sup>, Bassam B. Damaj<sup>5</sup>, Zivko L. Nikolov<sup>4</sup> & Kranthi K. Mandadi<sup>1,6</sup>✉

Plants represent a safe and cost-effective platform for producing high-value proteins with pharmaceutical properties; however, the ability to accumulate these in commercially viable quantities is challenging. Ideal crops to serve as biofactories would include low-input, fast-growing, high-biomass species such as sugarcane. The objective of this study was to develop an efficient expression system to enable large-scale production of high-value recombinant proteins in sugarcane culms. Bovine lysozyme (BvLz) is a potent broad-spectrum antimicrobial enzyme used in the food, cosmetics and agricultural industries. Here, we report a novel strategy to achieve high-level expression of recombinant proteins using a combinatorial stacked promoter system. We demonstrate this by co-expressing *BvLz* under the control of multiple constitutive and culm-regulated promoters on separate expression vectors and combinatorial plant transformation. *BvLz* accumulation reached 1.4% of total soluble protein (TSP) (10.0 mg *BvLz*/kg culm mass) in stacked multiple promoter:*BvLz* lines, compared to 0.07% of TSP (0.56 mg/kg) in single promoter:*BvLz* lines. *BvLz* accumulation was further boosted to 11.5% of TSP (82.5 mg/kg) through event stacking by re-transforming the stacked promoter:*BvLz* lines with additional *BvLz* expression vectors. The protein accumulation achieved with the combinatorial promoter stacking expression system was stable in multiple vegetative propagations, demonstrating the feasibility of using sugarcane as a biofactory for producing high-value proteins and bioproducts.

Recombinant proteins are currently being produced in cultured cell-based systems in mammals, microbes (bacteria and yeast), insects and plants, as well as in transgenic animals (reviewed by Demain and Vaishnav)<sup>1</sup>. Transgenic plants constitute an attractive system for expression and production of a variety of proteins and biomolecules due to their efficient eukaryotic protein synthesis, high scalability, relatively low production costs and environmental footprint<sup>2–4</sup>. However, selecting suitable hosts and expression vectors are key considerations since protein accumulation is determined by expression levels.

Important factors to consider when selecting a plant-based production platform include biomass yield per hectare, recombinant protein yield per unit biomass, ease of transformation, scalability and safety<sup>5</sup>. Sugarcane

<sup>1</sup>Texas A&M AgriLife Research and Extension Center, 2415 East US Highway 83, Weslaco, TX 78596, USA. <sup>2</sup>Department of Horticultural Sciences, Texas A&M University, College Station, TX 77843-2133, USA. <sup>3</sup>National Center for Therapeutics Manufacturing, Texas A&M University, 100 Discovery Drive, College Station, TX 77843-4482, USA. <sup>4</sup>BioSeparation Laboratory, Biological and Agricultural Engineering Department, College Station, TX 77843-2117, USA. <sup>5</sup>Innovus Pharmaceuticals, Inc., 8845 Rehco Road, San Diego, CA 92121, USA. <sup>6</sup>Department of Plant Pathology and Microbiology, Texas A&M University, College Station, TX 77843-2132, USA. <sup>7</sup>Present address: College of Medicine, Texas A&M University, 8447 Riverside Parkway, Bryan, TX 77807, USA. ✉email: mbdamaj@ag.tamu.edu; kkmmandadi@tamu.edu

(*Saccharum* spp. hybrids), a key feedstock in the expanding bioeconomy as a sugar and bioenergy crop<sup>6</sup>, is an ideal platform for recombinant protein production for several reasons: (1) It is a relatively fast growing tropical grass with the highly efficient  $C_4$  photosynthetic pathway, conferring high biomass production capacity with yields of up to 41.3 tons of biomass (harvested dry mass) per hectare per annum<sup>7,8</sup>; (2) it is highly efficient in utilizing radiation, water and nutrients to produce a large biomass and hence a higher recombinant protein yield; (3) it is readily amenable to genetic engineering, with established transformation and tissue regeneration techniques<sup>9,10</sup>; and (4) it has a low risk of out-crossing recombinant genes due to its primarily vegetative means of propagation; natural reproductive propagation in many temperate and subtropical regions is rare due to its photoperiod sensitivity.

Sugarcane was used as biofactory for the production of new biomolecules such as bioplastics<sup>11–15</sup>, alternative sugars (sorbitol and isomaltulose)<sup>16–18</sup>, and recombinant proteins including the human cytokine granulocyte macrophage colony stimulating factor GM-CSF<sup>19</sup>, canecystatins (cysteine protease inhibitors) CaneCP-1, CaneCP-2 and CaneCP-3<sup>20–22</sup>, and the cellulolytic enzymes, endoglucanase and cellobiohydrolases I and II<sup>23,24</sup>. Accumulation levels of these recombinant proteins ranged from 0.02 to 2.0% of total soluble protein (TSP) in leaves. However, very few attempts have so far been made to express recombinant proteins in sugarcane culms (reporter proteins)<sup>25</sup>, which constitute the largest fraction of harvestable biomass and would be an ideal platform for production of bulk proteins.

Bovine lysozyme (BvLz) is more important industrially than other lysozymes because of its potent broad-spectrum antimicrobial activity<sup>26,27</sup>, especially against Gram-negative bacteria and fungi at concentrations as low as 25 ppm, its sixfold higher chitinase activity than that of chicken lysozyme<sup>28</sup>, and its thermal stability and resistance to proteolysis<sup>29</sup>. BvLz, unlike other enzymes, possesses biochemical properties that make it suitable for protein extraction and purification, such as stability over a broad pH range, thermal stability, resistance to proteolysis and convenient quantification assays<sup>30,31</sup>.

In this study, we demonstrate the feasibility of developing sugarcane as an expression platform for production and purification of recombinant proteins at high levels, i.e. up to 11.5% of TSP (82.5 mg protein/kg culm mass). Multiple promoters (constitutive or culm-regulated) on separate expression vectors were stacked by combinatorial plant transformation approach to boost production levels of recombinant *bovine lysozyme* (BvLz), which was codon-optimized for expression in monocots. A double terminator or 3' untranslated region (UTR) was incorporated for improved transcript stability. Enzymatic activity and enzyme-linked immunosorbent assays (ELISA) of BvLz transgenic sugarcane culm protein extracts and clarified juice confirmed the presence of an intact and fully active BvLz enzyme, which accumulated in multiple vegetative generations at levels as high as 10.0 mg/kg (1.4% of TSP) in lines co-expressing BvLz from stacks of three or four different promoters on separate vectors, compared to 0.56 mg/kg (0.07% of TSP) in lines expressing BvLz from a single promoter vector. We further observed BvLz accumulation up to 82.5 mg/kg (11.5% of TSP) through event stacking by re-transforming the stacked promoter::BvLz transgenic lines with additional BvLz expression vectors.

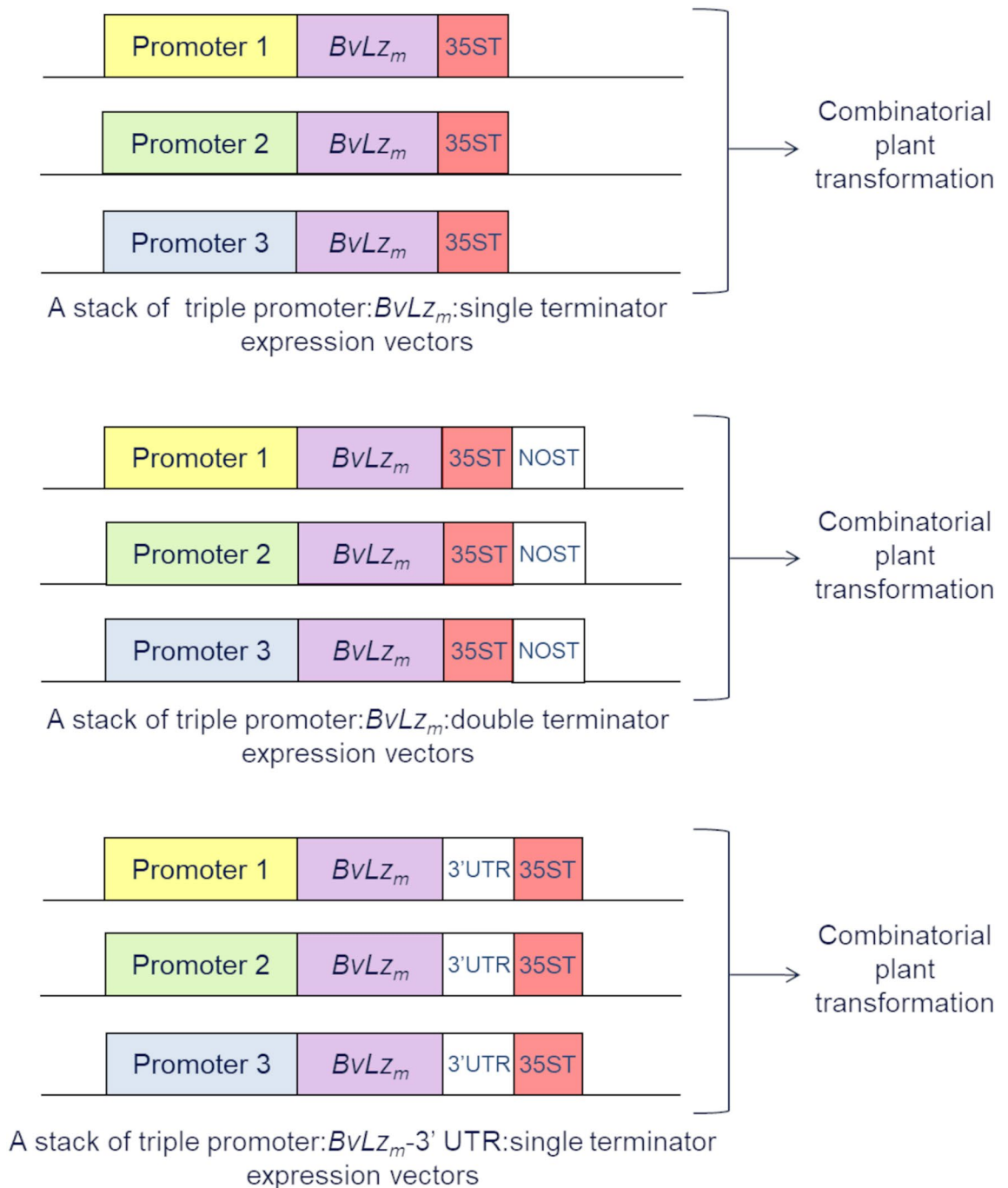
## Results and discussion

**The combinatorial promoter and event stacking result in increased recombinant protein production in transgenic sugarcane culms.** A salient feature of combinatorial transformation, a special case of co-transformation<sup>32</sup>, is that there is no theoretical limit to the number of expression vectors that can be co-transformed. To enable high-levels of recombinant protein production in sugarcane culms, we developed a combinatorial promoter and event stacking system and demonstrated its application in producing a high-value bovine lysozyme (BvLz) protein. This was facilitated by the availability of a set of constitutive and culm-regulated promoters previously isolated from sugarcane, in addition to the common maize *ubiquitin 1* promoter (pUbi)<sup>33</sup>. These include the culm-regulated promoters for *Sugarcane bacilliform virus* (pSCBV21)<sup>34</sup> and sugarcane *dirigent16* (pSHDIR16) gene<sup>35</sup>, and the constitutive promoters for sugarcane *proline-rich protein* (pSHPRP)<sup>36</sup> and *elongation factor 1 $\alpha$*  (pSHEF1 $\alpha$ )<sup>36</sup> genes. Furthermore, conditions for small-scale and large-scale extraction and clarification of recombinant BvLz from sugarcane culm extracts and juice were optimized at our Pilot Plant and BioSeparation Facilities<sup>30,37</sup>.

The essential design of the resulting new combinatorial promoter stacking system is illustrated in Fig. 1. The system consisted of co-expressing the codon-optimized BvLz<sub>m</sub>, from a stack of multiple promoters on separate expression vectors in sugarcane by combinatorial transformation. A double terminator, composed of the *Cauliflower mosaic virus* (CaMV) 35S terminator (35ST) and the *Agrobacterium tumefaciens* nopaline synthase terminator (NOST), or the 3'UTR of *Sorghum mosaic virus* (SrMV), was fused to the coding region of BvLz<sub>m</sub> to enhance transcript stability<sup>38,39</sup> (Fig. 1).

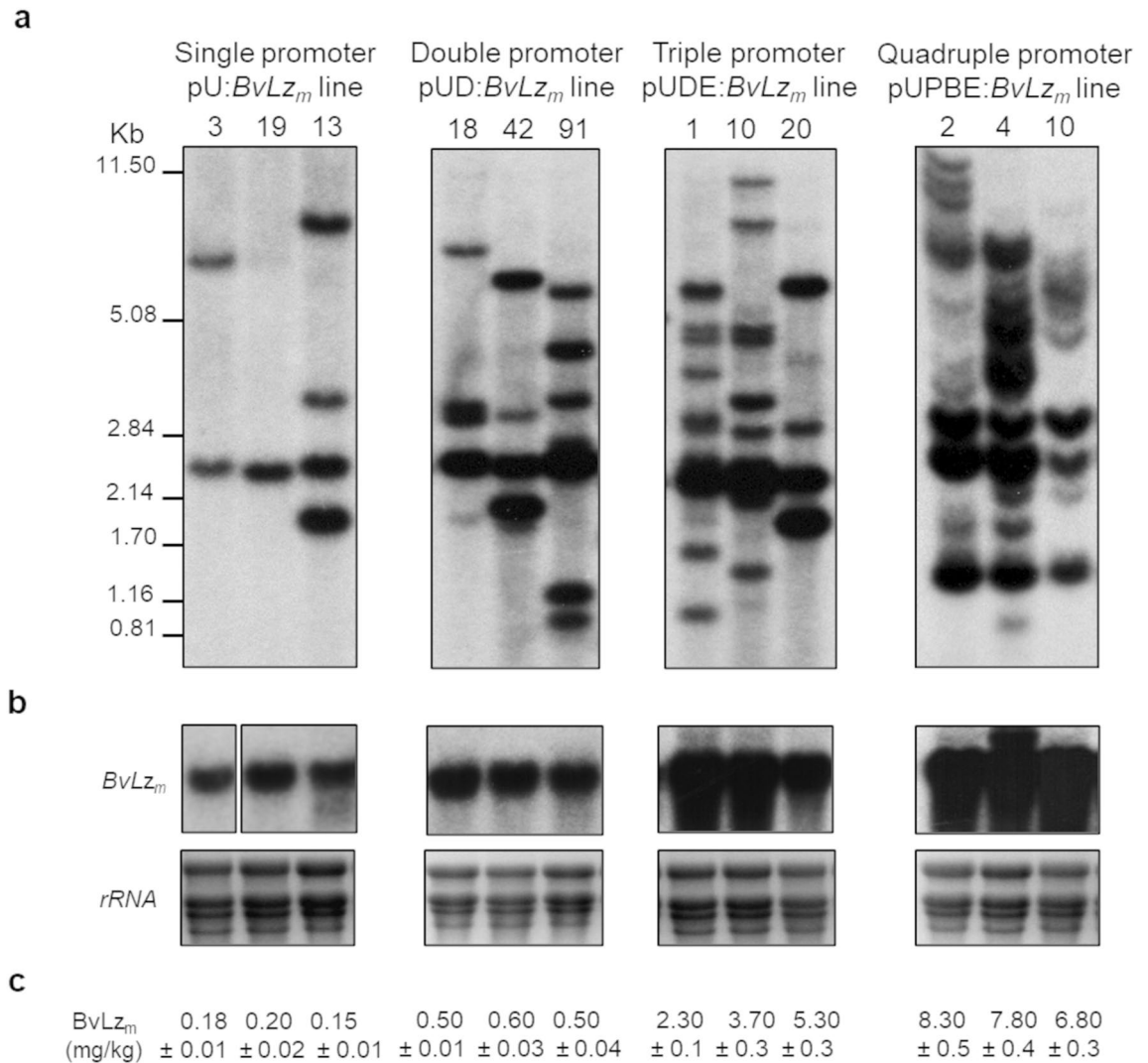
To test the stacking promoter gene expression system, embryogenic calli (2 month-old) and leaf roll discs (12 day-old), prepared from several commercial sugarcane varieties were co-transformed biolistically with the multiple promoter::BvLz<sub>m</sub> expression vectors, using the *bar* gene (phosphinothricin acetyl transferase) as a selectable marker. Several independent transgenic BvLz<sub>m</sub> lines, identified by Southern blot analysis (Fig. 2a; Supplementary Fig. S2), were generated from the combinatorial transformation of sugarcane with single, double, triple or quadruple promoter::BvLz<sub>m</sub> expression vectors (Table 1). These represent: (1) 43 lines (114 plants) expressing BvLz<sub>m</sub> from a single promoter, (2) 10 lines (52 plants) expressing BvLz<sub>m</sub> from a double promoter stack, (3) 24 lines (318 plants) expressing BvLz<sub>m</sub> from a triple promoter stack, and (4) 23 lines (76 plants) expressing BvLz<sub>m</sub> from a quadruple promoter stack (Table 1).

The integration and size of each respective BvLz<sub>m</sub> expression vector (promoter, BvLz<sub>m</sub>, terminator and/or 3'UTR) in the single and stacked multiple promoter::BvLz<sub>m</sub> lines were confirmed by Southern blot hybridization with a full-length BvLz<sub>m</sub> probe (Fig. 2a; Supplementary Fig. S2) and by PCR using primers encompassing each of the different promoter::BvLz<sub>m</sub>-terminator cassettes (Fig. 3; Supplementary Fig. S4, S5 and S6). All lines were



**Figure 1.** Design of a representative stacked multiple promoter:recombinant gene expression system developed for sugarcane. Promoter 1, 2 and 3 combinations can be any combination of the constitutive promoters maize, *ubiquitin 1*, sugarcane *proline rich protein* and sugarcane *elongation factor 1 $\alpha$*  or the culm-regulated promoters from sugarcane *dirigent16* and *Sugarcane bacilliform virus*. Vector assembly and cloning sites are indicated under “[Materials and methods](#)” section. *BvLz<sub>m</sub>*, maize codon-optimized *bovine lysozyme* gene; 35ST, terminator derived from *Cauliflower mosaic virus* 35S RNA; NOST, *Agrobacterium tumefaciens* nopaline synthase terminator; 3'UTR, 3' untranslated region of *Sorghum mosaic virus*.

analyzed for their *BvLz<sub>m</sub>* transcript levels by northern blot hybridization (Fig. 2b; Supplementary Fig. S3) as well as for their *BvLz<sub>m</sub>* accumulation by ELISA (Table 1; Fig. 2c for ELISA). For representative lines, yield was also determined by an enzyme activity assay and the results highly correlated with the ELISA data ( $R = 0.81\text{--}0.98$ ; Supplementary Table S1). Furthermore, in general, a clear positive trend was observed between the *BvLz<sub>m</sub>* copy number, the combinatorial promoter-*BvLz<sub>m</sub>* cassettes transformed and the *BvLz<sub>m</sub>* levels (Table 2; Fig. 2; Supplementary Fig. S2). For instance, quadruple and triple promoter:*BvLz<sub>m</sub>* lines displayed a higher *BvLz<sub>m</sub>* copy number and yield than double and single promoter:*BvLz<sub>m</sub>* lines, as expected from co-transformation (Table 2;



**Figure 2.** Stable integration, expression and yield of the *bovine lysozyme* (*BvLz<sub>m</sub>*) recombinant gene in sugarcane *BvLz<sub>m</sub>* transgenic lines as determined by Southern (a) and northern (b) blot analyses and enzyme-linked immunosorbent assay (ELISA) (c), respectively. Representative lines with single or multiple promoter:*BvLz<sub>m</sub>*-terminator cassettes are shown. *BvLz<sub>m</sub>*, maize codon-optimized *BvLz*; pU:*BvLz<sub>m</sub>*, *BvLz<sub>m</sub>* driven by the maize *ubiquitin 1* promoter (pU); pUD:*BvLz<sub>m</sub>*, *BvLz<sub>m</sub>* expressed from two promoters, pU and sugarcane *dirigent 16* (pD); pUDE:*BvLz<sub>m</sub>*, *BvLz<sub>m</sub>* expressed from three promoters, pU, pD and sugarcane *elongation factor 1α* (pE); and pUPBE:*BvLz<sub>m</sub>*, *BvLz<sub>m</sub>* expressed from four promoters, pU, sugarcane *proline-rich protein* (pP), *Sugarcane bacilliform virus* (pB) and pE. DNA and RNA gel blots were hybridized to a probe corresponding to the coding region of *BvLz<sub>m</sub>*. The full-length uncropped DNA and RNA gel blot autoradiograms are displayed in Supplementary Figures S2 and S3, respectively. The *BvLz<sub>m</sub>* yield is indicated as determined by ELISA in juice extract of culms (1.0 kg of culm).

Fig. 2; Supplementary Fig. S2). Similarly, the double promoter pUD:*BvLz<sub>m</sub>* lines had a higher *BvLz<sub>m</sub>* copy number and accumulation than single promoter pU:*BvLz<sub>m</sub>* lines (Table 2; Fig. 2; Supplementary Fig. S2).

The *BvLz<sub>m</sub>* yield from single promoter pUbi:*BvLz<sub>m</sub>* (pU:*BvLz<sub>m</sub>*) lines varied from low (0.08–0.1 mg/kg; 6.7% of plants) to moderate (0.12–0.18 mg/kg; 40.0% of plants) and high (0.2–0.4 mg/kg; 53.3% of plants) (Tables 1, 2). The *BvLz<sub>m</sub>* yield range was 0.08–0.4 mg/kg (0.01–0.06% of TSP), averaging 0.3 mg/kg (0.04% of TSP) ± 0.02 for the high expressers (Table 1). Other single promoter:*BvLz<sub>m</sub>* lines harboring pSHDIR16, pSCBV21, pSHPRP or pSHEF1α showed similar trends, with a highest *BvLz<sub>m</sub>* yield of 0.56 mg/kg (0.08% of TSP) (Supplementary Table S2).

Stacked double promoter pUbi-SHDIR16:*BvLz<sub>m</sub>* (pUD:*BvLz<sub>m</sub>*) lines displayed 1.8–6.3 fold higher *BvLz<sub>m</sub>* yield than single promoter pU:*BvLz<sub>m</sub>* lines, with a range of 0.5–0.7 mg/kg (0.07–0.1% of TSP) (Table 1). The *BvLz<sub>m</sub>* yield was further enhanced to 2.0–8.6 fold in the stacked triple promoter:*BvLz<sub>m</sub>* lines, with levels ranging from 1.0 to 6.0 mg/kg (0.1–0.8% of TSP) (Table 1). The majority (66.7%) of the stacked triple promoter pUbi-SHPRP-SHEF1α:*BvLz<sub>m</sub>* (pUPE:*BvLz<sub>m</sub>*) lines had a *BvLz<sub>m</sub>* yield of 1.0–2.0 mg/kg (0.1–0.3% of TSP) with 20.0% at 2.2–3.2 mg/kg (0.33–0.45% of TSP) and 13.3% at 3.5–4.7 mg/kg (0.5–0.7% of TSP) (Table 1). Replacing the constitutive SHPRP promoter with the culm-regulated SHDIR16 promoter in the stacked triple promoter

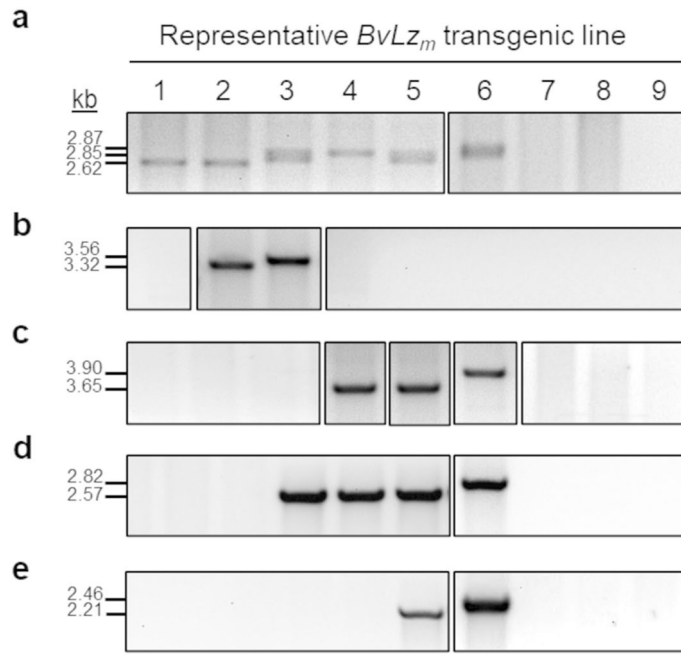
Bovine lysozyme (BvLz <sub>m</sub> ) expressing line and percentage of plants	BvLz <sub>m</sub> yield as determined by ELISA	
	BvLz <sub>m</sub> (mg/kg culm mass)	TSP (%)
<b>Promoter stacking</b>		
<i>1. Single promoter: BvLz<sub>m</sub> lines</i>		
pU: BvLz <sub>m</sub> ; single terminator (35ST) lines (43 lines; 114 plants)	0.08–0.4 (range)	0.01–0.06 (range)
6.70%	0.08–0.1	0.01–0.015
40.00%	0.12–0.18	0.02–0.027
53.30%	0.2–0.4	0.03–0.06
<i>2. Double promoter: BvLz<sub>m</sub> lines</i>		
pUD: BvLz <sub>m</sub> ; single terminator (35ST) lines (10 lines; 52 plants)	0.5–0.7 (range)	0.07–0.1 (range)
71.00%	0.5–0.58	0.07–0.077
29.00%	0.6–0.7	0.08–0.1
<i>3. Triple promoter: BvLz<sub>m</sub> lines</i>		
pUPE: BvLz <sub>m</sub> ; 3'UTR–single terminator (35ST) (10 lines; 32 plants)	1.0–4.7 (range)	0.1–0.7 (range)
66.70%	1.0–2.0	0.1–0.3
20.00%	2.2–3.2	0.33–0.45
13.30%	3.5–4.7	0.5–0.7
pUDE: BvLz <sub>m</sub> ; 3'UTR–single terminator (35ST) (14 lines; 286 plants)	1.5–6.0 (range)	0.2–0.8 (range)
27.00%	1.5–2.0	0.2–0.3
62.00%	2.2–3.2	0.33–0.45
4.50%	3.5–4.7	0.5–0.7
6.50%	5.0–6.0	0.7–0.8
<i>4. Quadruple promoter: BvLz<sub>m</sub> lines</i>		
pUPBE: BvLz <sub>m</sub> ; 3'UTR–single terminator (35ST) (12 lines; 36 plants)	2.0–6.3 (range)	0.3–0.9 (range)
33.30%	2.0–3.5	0.3–0.5
22.20%	4.0–5.5	0.6–0.77
44.50%	6.0–6.3	0.8–0.9
pUPBE: BvLz <sub>m</sub> ; double terminator (35STNOST) (11 lines; 40 plants)	2.4–10.0 (range)	0.3–1.4 (range)
24.20%	2.4–3.5	0.3–0.5
51.70%	4.0–5.5	0.6–0.77
24.10%	6.0–10.0	0.8–1.4
<b>Event stacking</b>		
<i>Five promoter: BvLz<sub>m</sub> lines</i>		
pUDE: BvLz <sub>m</sub> line + pP: BvLz <sub>m</sub> + pB: BvLz <sub>m</sub> (44 lines; 110 plants)	11.0–82.5 (range)	1.5–11.5 (range)
12.10%	11.0–12.4	1.5–1.7
24.20%	15.9–21.1	2.2–2.9
33.30%	26.2–32.3	3.6–4.5
18.30%	59.9–82.5	8.3–11.5

**Table 1.** Recombinant bovine lysozyme yield of transgenic sugarcane culm. The percentage (%) of lines with the corresponding BvLz<sub>m</sub> yield are indicated for each single or stacked promoter: BvLz<sub>m</sub> construct. BvLz<sub>m</sub>, maize codon-optimized bovine lysozyme gene; U, maize ubiquitin 1 promoter; D, sugarcane dirigent16 promoter; P, sugarcane proline rich protein promoter; E, sugarcane elongation factor 1 $\alpha$  promoter; B, *Sugarcane bacilliform virus* promoter; 3'UTR, 3' untranslated region of *Sorghum mosaic virus*; 35ST, *Cauliflower mosaic virus* 35S terminator; NOST, *Agrobacterium tumefaciens* nopaline synthase terminator; ELISA, enzyme-linked immunosorbent assay; TSP, total soluble protein.

pUbi-SHDIR16-SHEF1 $\alpha$ :BvLz<sub>m</sub> (pUDE: BvLz<sub>m</sub>) lines boosted the BvLz<sub>m</sub> yield to 6.0 mg/kg (0.8% of TSP). Most of pUDE: BvLz<sub>m</sub> lines (62.0%) had a BvLz<sub>m</sub> yield of 2.2–3.2 mg/kg (0.33–0.45% of TSP), with 27.0% at 1.5–2.0 mg/kg (0.2–0.3% of TSP), 4.5% at 3.5–4.7 mg/kg (0.5–0.7% of TSP) and 6.5% at 5.0–6.0 mg/kg (0.7–0.8% of TSP) (Table 1).

Next, we checked if stacking another promoter to produce quadruple promoter: BvLz<sub>m</sub> lines would be helpful. The BvLz<sub>m</sub> yield increased modestly in the stacked quadruple promoter pUbi-SHPRP-SCBV21-SHEF1 $\alpha$ : BvLz<sub>m</sub> (pUPBE: BvLz<sub>m</sub>) lines by 1.7–2.4 fold, compared to the stacked triple promoter: BvLz<sub>m</sub> lines. The highest enhancement was achieved when using a double terminator cassette, i.e. 10.0 mg/kg (1.4% of TSP) in pUPBE: BvLz<sub>m</sub>; 35STNOST lines (Table 1), and the 3'UTR of SrMV with the single 35S terminator, i.e. 6.3 mg/kg (0.9% of TSP) in pUPBE: BvLz<sub>m</sub>; 3'UTR35ST lines (Table 1). In fact, 24.1% of pUPBE: BvLz<sub>m</sub>; 35STNOST plants had a BvLz<sub>m</sub> yield of 6.0–10 mg/kg (0.8–1.4% of TSP), and 44.5% of pUPBE: BvLz<sub>m</sub>; 3'UTR35ST plants showed a BvLz<sub>m</sub> yield of 6.0–6.3 mg/kg (0.8–0.9% of TSP) (Table 1). Lastly, we evaluated if event stacking can enhance the





**Figure 3.** Presence and size of multiple promoter:*bovine lysozyme* (*BvLz<sub>m</sub>*)-terminator cassettes in the same *BvLz<sub>m</sub>* transgenic line as determined by PCR analysis. Representative lines with single or multiple promoter:*BvLz<sub>m</sub>*-terminator cassettes are shown. (1) pU:*BvLz<sub>m</sub>*-35ST line; (2) pUD:*BvLz<sub>m</sub>*-35ST line; (3) pUDE:*BvLz<sub>m</sub>*-3'UTR-35ST line; (4) pUPE:*BvLz<sub>m</sub>*-3'UTR-35ST line; (5) pUPBE:*BvLz<sub>m</sub>*-3'UTR-35ST line; (6) pUPBE:*BvLz<sub>m</sub>*-35STNOST line; (7) vector-transformed line; (8) non-transformed (NT; tissue culture-derived) plant; and (9) no DNA template (negative control for PCR). (a) Detection of pUbi, *BvLz<sub>m</sub>*, 3'UTR, 35ST and NOST using the primer sets pUbi-F/35ST-R (2.62 kilobase pairs [kb] or 2.85 kb fragment) and pUbi-F/NOST-R (2.87 kb fragment). (b) Detection of pSHDIR16, *BvLz<sub>m</sub>*, 3'UTR, 35ST and NOST using the primer sets pSHDIR16-F/35ST-R (3.32 kb fragment) and pSHDIR16-F/NOST-R (3.56 kb fragment). (c) Detection of pSHPRP, *BvLz<sub>m</sub>*, 3'UTR, 35ST and NOST using the primer sets pSHPRP-F/35ST-R (3.65 kb fragment) and pSHPRP-F/NOST-R (3.90 kb fragment). (d) Detection of pSHEF1α, *BvLz<sub>m</sub>*, 3'UTR, 35ST and NOST using the primer sets pSHEF1α-F/35ST-R (2.57 kb fragment) and pSHEF1α-F/NOST-R (2.82 kb fragment). (e) Detection of pSCBV21, *BvLz<sub>m</sub>*, 3'UTR, 35ST and NOST using the primer sets pSCBV21-F/35ST-R (2.21 kb fragment) and pSCBV21-F/NOST-R (2.46 kb fragment). *BvLz<sub>m</sub>*, maize codon-optimized *bovine lysozyme* gene; U, Ubi promoter; D, SHDIR16 promoter; P, SHPRP promoter; E, SHEF1α promoter; B, SCBV21 promoter; 3'UTR, 3' untranslated region of *Sorghum mosaic virus*. 35ST, *Cauliflower mosaic virus* 35S terminator; NOST, *Agrobacterium tumefaciens* nopaline synthase terminator. Full-length uncropped gels of the PCR products are displayed in Supplementary Figures S4, S5 and S6.

yields of the stacked quadruple promoter lines. Event stacking, also referred to as super transformation, is a good alternative to hybridization/crossing, which is time-consuming and not a viable option in vegetatively-propagated crops like sugarcane. Stacked five promoter pUbi-SHDIR16-SHEF1α-SHPRP-SCBV21:*BvLz<sub>m</sub>* (pUDEPB:*BvLz<sub>m</sub>*) lines were generated through event stacking, by re-transforming bialaphos-resistant triple promoter pUDE:*BvLz<sub>m</sub>* lines with two promoter:*BvLz<sub>m</sub>* expression vectors, pP:*BvLz<sub>m</sub>* and pB:*BvLz<sub>m</sub>* (Table 1) using the *neomycin phosphotransferase II* as a selectable marker. The resulting pUDEPB:*BvLz<sub>m</sub>* lines showed increased *BvLz<sub>m</sub>* accumulation, i.e. up to 82.5 mg/kg culm mass (11.5% of TSP) (Table 1). The majority (33.3%) of these lines exhibited *BvLz<sub>m</sub>* levels of 26.2–32.3 mg/kg (3.6–4.5% of TSP), while 18.3% accumulated the highest *BvLz<sub>m</sub>* levels, i.e. 59.9–82.5 mg/kg (8.3–11.5% of TSP). The remaining 24.2% and 12.1% of the lines showed *BvLz<sub>m</sub>* levels of 15.9–21.1 mg/kg (2.2–2.9% of TSP) and 11.0–12.4 mg/kg (1.5–1.7% of TSP), respectively (Table 1). Notably, *BvLz<sub>m</sub>* accumulation was highly enhanced in the new stacked five promoter pUDEPB:*BvLz<sub>m</sub>* lines by 7.3–13.8-fold, compared to the receiving stacked triple promoter pUDE:*BvLz<sub>m</sub>* lines. Together, these experiments demonstrate that high levels of recombinant *BvLz<sub>m</sub>* (up to 11.5% of TSP or 82.5 mg/kg) can be successfully produced in sugarcane culms using the combinatorial promoter and event stacking strategies. Previous studies utilized multiple plant species, tissue types, and expression systems for recombinant protein production<sup>40,41</sup>. Majority of them used transient *Agrobacterium*- and viral vector-based approaches in *Nicotiana benthamiana* or *N. tabacum*<sup>42–46</sup>. While the transient systems are viable approaches, they are technically feasible only in few plant species that are amenable for infiltration and/or are hosts for the viruses used as viral vectors. In this context, transgenic plant systems are more suited for wider adoption since broad range of plant species can be transformed using latest biotechnology tools. When comparing our results of protein expression in sugarcane culms with other transgenic plant expression systems, caution was exercised particularly when comparing recovered protein yields per starting tissue weight (e.g., mg/kg). This is because not all plant tissues have similar compositions, nor the protein extractions are equally efficient

Bovine lysozyme ( <i>BvLz<sub>m</sub></i> ) expressing line	<i>BvLz<sub>m</sub></i> yield (mg/kg culm mass)	Average <i>BvLz<sub>m</sub></i> copy number				
		pU- <i>BvLz<sub>m</sub></i> junction	pD- <i>BvLz<sub>m</sub></i> junction	pP- <i>BvLz<sub>m</sub></i> junction	pE- <i>BvLz<sub>m</sub></i> junction	pB- <i>BvLz<sub>m</sub></i> junction
pU: <i>BvLz<sub>m</sub></i> -single terminator (35ST)						
3	0.18 ± 0.01	4.0 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
13	0.20 ± 0.02	1.0 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
19	0.15 ± 0.01	2.0 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
pUD: <i>BvLz<sub>m</sub></i> -single terminator (35ST)						
18	0.50 ± 0.01	6.0 ± 0.3	4.0 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
42	0.60 ± 0.03	7.0 ± 0.3	5.0 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
91	0.50 ± 0.04	6.0 ± 0.2	7.0 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
pUPE: <i>BvLz<sub>m</sub></i> -3'UTR-35ST						
11	1.4 ± 0.1	6.0 ± 0.3	0.0 ± 0.0	18.0 ± 4.1	6.0 ± 0.2	0.0 ± 0.0
22	2.0 ± 0.2	6.0 ± 0.2	0.0 ± 0.0	15.0 ± 0.5	3.0 ± 0.4	0.0 ± 0.0
24	3.5 ± 0.3	5.0 ± 0.1	0.0 ± 0.0	11.0 ± 0.6	4.0 ± 0.4	0.0 ± 0.0
pUDE: <i>BvLz<sub>m</sub></i> -3'UTR-35ST						
1	2.3 ± 0.1	8.0 ± 0.8	13.0 ± 1.5	0.0 ± 0.0	6.0 ± 0.5	0.0 ± 0.0
10	3.7 ± 0.3	6.0 ± 0.4	10.0 ± 0.9	0.0 ± 0.0	5.0 ± 0.1	0.0 ± 0.0
20	5.3 ± 0.3	10.0 ± 1.0	8.0 ± 0.6	0.0 ± 0.0	3.0 ± 0.3	0.0 ± 0.0
pUPBE: <i>BvLz<sub>m</sub></i> -3'UTR-35ST						
1	6.3 ± 0.4	10.0 ± 0.9	0.0 ± 0.0	17.0 ± 1.6	6.0 ± 0.3	15.0 ± 1.2
4	6.0 ± 0.3	5.0 ± 0.3	0.0 ± 0.0	20.0 ± 2.6	4.0 ± 0.1	6.0 ± 0.4
15	6.0 ± 0.2	6.0 ± 0.4	0.0 ± 0.0	12.0 ± 1.4	8.0 ± 0.4	11.0 ± 0.7
pUPBE: <i>BvLz<sub>m</sub></i> -35STNOST						
1	6.7 ± 0.3	6.0 ± 0.3	0.0 ± 0.0	30.0 ± 1.4	7.0 ± 0.8	18.0 ± 1.6
2	10.0 ± 0.7	3.0 ± 0.2	0.0 ± 0.0	11.0 ± 0.4	3.0 ± 0.2	7.0 ± 0.7
4	8.3 ± 0.4	11.0 ± 0.9	0.0 ± 0.0	25.0 ± 1.6	5.0 ± 0.3	25.0 ± 2.6
pUDE: <i>BvLz<sub>m</sub></i> + pP: <i>BvLz<sub>m</sub></i> + pB: <i>BvLz<sub>m</sub></i>						
1	11.8 ± 1.4	10.0 ± 1.3	23.0 ± 3.0	15.0 ± 1.6	8.0 ± 0.9	10.0 ± 1.0
5	14.6 ± 2.3	8.0 ± 1.2	13.0 ± 1.4	25.0 ± 2.7	10.0 ± 1.5	22.0 ± 2.4
12	28.6 ± 3.4	14.0 ± 1.5	17.0 ± 1.9	21 ± 2.3	12.0 ± 1.9	31.0 ± 4.1

**Table 2.** Average gene copy number for representative *bovine lysozyme* expressing lines as determined by quantitative PCR. The gene copy number was estimated based on gene copy number indices generated using the reference gene prolyl 4-hydroxylase. *BvLz<sub>m</sub>*, maize codon-optimized *BvLz*; pU, maize *ubiquitin 1* promoter; pUD:*BvLz<sub>m</sub>*, *BvLz<sub>m</sub>* expressed from two promoters, maize *ubiquitin 1* and sugarcane *dirigent16* (pD); pUPE:*BvLz<sub>m</sub>*, *BvLz<sub>m</sub>* expressed from three promoters, pU, sugarcane *proline-rich protein* (pP) and sugarcane *elongation factor 1α* (pE); pUDE:*BvLz<sub>m</sub>*, *BvLz<sub>m</sub>* expressed from three promoters, pU, pD and pE; pUPBE:*BvLz<sub>m</sub>*, *BvLz<sub>m</sub>* expressed from four promoters, pU, pP, pE and *Sugarcane bacilliform virus*; 3'UTR, 3' untranslated region of *Sorghum mosaic virus*; 35ST, *Cauliflower mosaic virus* 35S terminator; NOST, *Agrobacterium tumefaciens* nopaline synthase terminator. The *BvLz<sub>m</sub>* yield is indicated as determined by enzyme-linked immunosorbent assay in juice extract of culms (one kg of culm).

among tissue types, owing to biological and biochemical differences<sup>41</sup>. For instance, sugarcane culms primarily constitute juice (sugars) and lignocellulosic fiber (bagasse). An equal amount of *N. benthamiana* leaves on a fresh weight basis will have less fiber, and proteins may be easier to extract from leaf tissues. We also note that biochemical properties of target proteins such as size, solubility, amino-acid composition, structural features, and protein stability may also ultimately influence the final yield. With these caveats in mind, we compared our results with other reported studies of transgenic plant systems using the % TSP unit of recovered proteins. Several studies have reported recombinant protein yields of ~0.002 to 0.05% of TSP in transgenic carrots<sup>47,48</sup>, ~0.23–2.5% of TSP in transgenic tobacco and potato<sup>49</sup>, ~8% TSP in transgenic tomato<sup>50</sup>, and ~11.9% in transgenic rice<sup>51</sup>. These comparisons suggest that higher protein yields can be achieved using the sugarcane transgenic system (up to 11.5% of TSP), which are comparable to other transgenic systems, if not greater.

In addition to the use of constitutive or tissue-specific promoters, inducible promoters can be used for expressing recombinant proteins in plants<sup>52–54</sup>. Several inducible promoters can be used for generating transgenic plants such as dexamethasone-, ethylene-, heat shock- and estradiol-inducible promoters<sup>52</sup>. Indeed, we have previously shown that the sugarcane DIRIGENT (SHDIR16) promoter is responsive to plant hormones such as salicylic acid or jasmonic acid<sup>35</sup>. This is promising and suggests that inducible promoters such as SHDIR16, and other well-characterized plant inducible-promoters<sup>52</sup> can be further used in lieu or in combination with the constitutive/tissue-specific promoters that we have described, in order to robustly control and/or fine-tune the recombinant protein expression.

**Increased protein levels were associated with the number of combinatorial stacked promoters and not with the copy number alone.** Our results show that using multiple different promoters to drive expression of recombinant *BvLz<sub>m</sub>* on distinct vectors enhanced recombinant protein accumulation. It is possible that the enhanced levels may have occurred due to higher number of inserted *BvLz<sub>m</sub>* copies alone or it could be due to a combination of promoter-driven synergistic transcriptional activity. To test these scenarios, we performed a comparison of the *BvLz<sub>m</sub>* transcript and yield among the various promoter stacked lines that had similar number of insertions. This analysis showed that there is a positive correlation in *BvLz<sub>m</sub>* transcript and yield with combinatorial promoter:*BvLz<sub>m</sub>* stacks, irrespective of the number *BvLz<sub>m</sub>* inserts (Fig. 2; Supplementary Fig. S2). For instance, for single promoter:*BvLz<sub>m</sub>* line 13, double promoter:*BvLz<sub>m</sub>* line 42, triple promoter:*BvLz<sub>m</sub>* line 20 and quadruple promoter:*BvLz<sub>m</sub>* line 10, with all of them having about 4–5 *BvLz<sub>m</sub>* inserts, there was a clear enhancement in the *BvLz<sub>m</sub>* yield (Fig. 2a,c; Supplementary Fig. S2). Conversely, a comparison of single promoter:*BvLz<sub>m</sub>* transgenic lines with one or multiple inserts showed that there was no corresponding increase in *BvLz<sub>m</sub>* yield with the copy number. For instance, line 19 with one insert (Fig. 2a,c; Supplementary Fig. S2) had a *BvLz<sub>m</sub>* yield of 0.2 mg/kg, while line 13 with 4 *BvLz<sub>m</sub>* inserts had a *BvLz<sub>m</sub>* yield of 0.15 mg/kg (Fig. 2a, c; Supplementary Fig. S2). Together, these results suggest that the increase in *BvLz<sub>m</sub>* yield is primarily attributed to the number of combinatorial stacked multiple promoters and not just with the *BvLz<sub>m</sub>* copy number alone.

**Combinatorial promoter stacking may alleviate transcriptional occlusion and/or recombinant gene silencing.** Multiple identical copies of recombinant genes or promoter transcription units (PTUs) delivered through a single construct could trigger transgene silencing<sup>55–57</sup> or result in promoter occlusion or transcriptional interference, a phenomenon observed in eukaryotic systems, including plants<sup>58–62</sup>. For instance, a strong PTU can sequester most of the transcription factors in its immediate vicinity, limiting transcription from other promoters present in *cis* on the same vector<sup>63</sup>. Alternatively, homology-dependent DNA methylation within the promoter or in the coding region sequences could result in transgene silencing. For instance, in maize, transgenic lines with four copies of a cellulase gene, under control of tandemly arranged PTUs on the same vector, resulted in lowered expression than those lines with fewer copies<sup>64</sup>.

Our results here showed a positive correlation between the number of combinatorial promoter stacks of recombinant *BvLz<sub>m</sub>* and increase in *BvLz<sub>m</sub>* levels, with no apparent transgene silencing. It is likely that using different promoter sequences in separate vectors may overcome the transgene silencing or transcriptional interference. We suggest that each expression vector in the described stacked multiple promoter:*BvLz<sub>m</sub>* system (Fig. 1) does not negatively affect the others, as shown by a positive correlation between the combinatorial promoter:*BvLz<sub>m</sub>* copy number (Table 2) and enhanced steady-state *BvLz<sub>m</sub>* transcript accumulation (Fig. 2b; Supplementary Fig. S3) and *BvLz<sub>m</sub>* activity (Table 1; Fig. 2c).

**Elevated recombinant *BvLz<sub>m</sub>* accumulation positively enhances transgenic plant growth.** Analysis of the deleterious effects of recombinant protein accumulation on plant physiology and growth is crucial in order to assess the economic feasibility of using transgenic plants as biofactories, and this is largely dependent on the target protein function<sup>65</sup>. In our scenario with *BvLz<sub>m</sub>*, we found no deleterious effects of enhanced *BvLz<sub>m</sub>* expression on sugarcane growth. On the contrary, several growth characteristics of *BvLz<sub>m</sub>* expressing lines were better than those of non-transformed plants, such as enhanced leaf length, culm height, tiller number, culm biomass and Brix (total soluble solids) (Table 3). These differences were statistically significant ( $p < 0.001$  and  $p < 0.0001$ ) in the triple and quadruple promoter:*BvLz<sub>m</sub>* expressing lines (Table 3). For instance, mean culm fresh biomass per plant of the quadruple promoter:*BvLz<sub>m</sub>* expressing lines was nearly 2.5 times greater than that of non-transformed plants. The mean soluble solids content in juice from triple and quadruple promoter:*BvLz<sub>m</sub>* expressing lines was approximately 20% higher than that of non-transformed plants. Similar trends were also observed for leaf length, culm height and tiller density. The enhanced agronomic performance of the transgenic lines suggested that *BvLz<sub>m</sub>*, which is a well-known antimicrobial protein<sup>27</sup>, could have a growth-promoting or perhaps protective role against pathogens present in the natural growth environment.

**Recombinant protein accumulation in culms increases with plant age.** To monitor the temporal stability of *BvLz<sub>m</sub>* accumulation in sugarcane culms in a growing season, we analyzed *BvLz<sub>m</sub>* levels for 11-months with a selection of several representative single promoter pU:*BvLz<sub>m</sub>* lines. The *BvLz<sub>m</sub>* yields (mg of *BvLz<sub>m</sub>*/kg of harvested culm) in these lines after 7-, 9- and 11-month-harvest are shown in Fig. 4 (data for four representative lines) and Table 3 (data for six representative lines at the 11 month-harvest). There was a significant ( $p < 0.05$ ) increase in *BvLz<sub>m</sub>* yield over time for all the lines tested. *BvLz<sub>m</sub>* accumulation was highest at the 11-month harvest, with lines 67, 108 and 114 showing the most significant ( $p < 0.05$ ) increase (Fig. 4). This accumulation pattern coincides with timing of culm ripening, which is characterized by increased sucrose translocation and accumulation in culms. The age-related sucrose accumulation also was associated with the reduction in vegetative development (leaf initiation and expansion) and commences at the mature basal internodes, progressing towards the culm apex, until the entire culm reaches a stable sugar level as it approaches physiological maturity<sup>66</sup>. The age-related pattern of *BvLz<sub>m</sub>* accumulation may also be regulated by similar factors whereby photoassimilates and other substrates for *BvLz<sub>m</sub>* are diverted from vegetative growth towards metabolite synthesis and accumulation during sugarcane maturation. Regardless of the mechanisms regulating the temporal accumulation of *BvLz<sub>m</sub>*, our results demonstrate that the recombinant protein levels can be maintained, if not enhanced, during the development phases of sugarcane in a growing season. Similar results were observed for *BvLz<sub>m</sub>* accumulation in representative triple promoter:*BvLz<sub>m</sub>* lines, which showed sustained and stable *BvLz<sub>m</sub>* levels over a full growing year, as well as in successive vegetative propagations (Supplementary Table S3). Similar accumulation of recombinant proteins (human therapeutic interleukin-10) with plant maturity was observed in tobacco<sup>67</sup>.

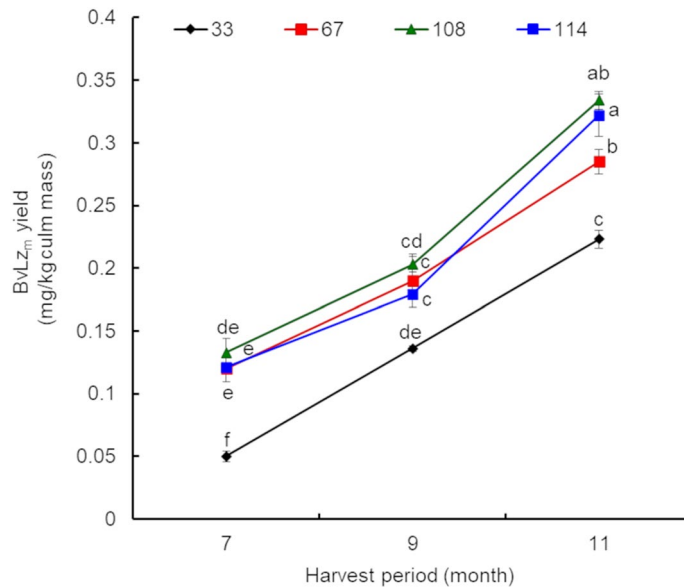


Bovine lysozyme ( <i>BvLz<sub>m</sub></i> ) expressing line	<i>BvLz<sub>m</sub></i> yield (mg/kg culm fresh mass)	Agronomic parameter				
		Leaf length (cm)	Culm height (cm)	Tiller number	Culm biomass (fresh mass, kg/ plant)	Brix (%)
Non-transformed (CP72-1210)	0.0 ± 0.0 <sup>a</sup>	75.3 ± 1.4 <sup>a</sup>	12.5 ± 0.2 <sup>a</sup>	4.0 ± 0.2 <sup>a</sup>	7.1 ± 0.3 <sup>a</sup>	14.6 ± 0.6 <sup>a</sup>
<b>Single promoter maizeubiquitin 1:<i>BvLz<sub>m</sub></i> expressing lines (CP72-1210)</b>						
33	0.22 ± 0.01 <sup>b*</sup>	74.0 ± 0.5 <sup>a</sup>	13.5 ± 0.2 <sup>a</sup>	10.0 ± 0.4 <sup>b**</sup>	9.1 ± 0.6 <sup>b*</sup>	17.2 ± 0.6 <sup>b*</sup>
67	0.27 ± 0.01 <sup>b*</sup>	72.3 ± 1.1 <sup>a</sup>	14.5 ± 0.6 <sup>a</sup>	4.0 ± 0.3 <sup>a</sup>	7.8 ± 0.5 <sup>a</sup>	15.7 ± 0.5 <sup>b*</sup>
108	0.33 ± 0.01 <sup>b*</sup>	76.8 ± 1.4 <sup>a</sup>	18.0 ± 0.3 <sup>b**</sup>	5.0 ± 0.2 <sup>a</sup>	9.6 ± 0.9 <sup>b*</sup>	14.9 ± 0.4 <sup>a</sup>
114	0.32 ± 0.01 <sup>b*</sup>	70.5 ± 2.2 <sup>a</sup>	11.3 ± 0.6 <sup>a</sup>	6.0 ± 0.2 <sup>b*</sup>	8.2 ± 0.6 <sup>b*</sup>	16.8 ± 0.4 <sup>b*</sup>
116	0.36 ± 0.01 <sup>b*</sup>	74.0 ± 3.3 <sup>a</sup>	14.5 ± 1.0 <sup>a</sup>	8.0 ± 0.9 <sup>b**</sup>	7.0 ± 0.3 <sup>a</sup>	17.9 ± 0.7 <sup>b*</sup>
123	0.26 ± 0.01 <sup>b*</sup>	76.5 ± 1.8 <sup>a</sup>	17.1 ± 1.3 <sup>b*</sup>	4.0 ± 0.1 <sup>a</sup>	8.9 ± 0.5 <sup>b*</sup>	15.2 ± 0.5 <sup>a</sup>
<b>Triple promoter:<i>BvLz<sub>m</sub></i> expressing lines</b>						
pUPE: <i>BvLz<sub>m</sub></i> 32 (CP72-1210)	2.7 ± 0.1 <sup>b**</sup>	81.9 ± 1.2 <sup>a</sup>	16.3 ± 0.4 <sup>b*</sup>	17.0 ± 2.9 <sup>b**</sup>	13.1 ± 1.2 <sup>b*</sup>	18.4 ± 0.3 <sup>b*</sup>
pUDE: <i>BvLz<sub>m</sub></i> (TCP98-4454)						
18	5.1 ± 0.3 <sup>b**</sup>	92.3 ± 4.1 <sup>b**</sup>	17.3 ± 0.2 <sup>b*</sup>	7.0 ± 0.6 <sup>b*</sup>	10.5 ± 0.9 <sup>b*</sup>	16.7 ± 0.3 <sup>b*</sup>
19	4.6 ± 0.2 <sup>b**</sup>	95.6 ± 2.5 <sup>b**</sup>	18.5 ± 0.4 <sup>b**</sup>	7.0 ± 0.2 <sup>b*</sup>	11.3 ± 1.0 <sup>b*</sup>	15.6 ± 0.4 <sup>a</sup>
44	2.9 ± 0.1 <sup>b**</sup>	74.8 ± 2.6 <sup>a</sup>	14.3 ± 1.4 <sup>a</sup>	7.0 ± 1.3 <sup>b*</sup>	8.6 ± 0.8 <sup>b*</sup>	16.0 ± 0.2 <sup>b*</sup>
54	6.0 ± 0.3 <sup>b**</sup>	87.2 ± 4.0 <sup>b*</sup>	16.6 ± 1.1 <sup>b*</sup>	4.0 ± 0.5 <sup>a</sup>	10.7 ± 1.1 <sup>b*</sup>	17.9 ± 0.6 <sup>b*</sup>
<b>Quadruple promoter:<i>BvLz<sub>m</sub></i> expressing lines</b>						
pUPBE: <i>BvLz<sub>m</sub></i> (TCP98-4454)						
1	6.7 ± 0.3 <sup>b**</sup>	115.3 ± 3.1 <sup>b**</sup>	19.4 ± 1.2 <sup>b**</sup>	21.1 ± 1.3 <sup>b**</sup>	16.4 ± 0.7 <sup>b**</sup>	19.1 ± 0.6 <sup>b*</sup>
2	10.0 ± 0.7 <sup>b**</sup>	201.8 ± 6.4 <sup>b**</sup>	32.1 ± 2.4 <sup>b**</sup>	17.0 ± 1.1 <sup>b**</sup>	20.8 ± 1.0 <sup>b**</sup>	18.0 ± 0.7 <sup>b*</sup>
4	8.3 ± 0.4 <sup>b**</sup>	140.2 ± 5.1 <sup>b**</sup>	25.3 ± 1.9 <sup>b**</sup>	14.0 ± 0.9 <sup>b**</sup>	18.6 ± 0.9 <sup>b**</sup>	17.5 ± 0.3 <sup>b*</sup>

**Table 3.** Growth and culm quality characteristics of bovine lysozyme expressing sugarcane lines. Morphological parameters (leaf height, culm height and tiller number) of 15 representative single promoter:bovine lysozyme (*BvLz<sub>m</sub>*) and five representative triple promoter:*BvLz<sub>m</sub>* expressing lines were measured every 2 weeks for 4 months after planting. Culm biomass was determined at 11 months (harvest). The *BvLz<sub>m</sub>* yield is indicated as determined by enzyme-linked immunosorbent assay in juice extract of culms at the 11-month harvest (one kg of culm). Total soluble solids (Brix) of extracted culm juice from 11-month-old plants was determined using a refractometer (model PR-101a, Atago U.S.A, Inc., Bellevue, WA). Data represent means from four biological replications ± standard errors. Means are compared column-wise. Values that are significantly different from those of non-transformed at  $p < 0.001$  and  $p < 0.0001$  are denoted by \* and \*\*, respectively. *BvLz<sub>m</sub>*: maize codon-optimized *BvLz*; pUPE:*BvLz<sub>m</sub>*: *BvLz<sub>m</sub>* expressed from three promoters, maize ubiquitin 1 (pUbi), sugarcane proline-rich protein (pSHPRP) and sugarcane elongation factor 1a (pSHEF1a); pUDE:*BvLz<sub>m</sub>*: *BvLz<sub>m</sub>* expressed from three promoters, pUbi, sugarcane dirigent16 and pSHEF1a. pUPBE:*BvLz<sub>m</sub>*: *BvLz<sub>m</sub>* expressed from four promoters, pUbi, pSHPRP, pSHEF1a and Sugarcane bacilliform virus promoter.

**High level recombinant protein accumulation requires adequate mineral nutrition to sustain the protein and biomass accumulation.** Adequate water and nutrients supply are important for crop productivity as well as quality considerations, such as protein content and other sensory traits<sup>68</sup>. Because we observed enhanced growth traits such as biomass in the *BvLz<sub>m</sub>* expressing lines, specifically in the triple promoter:*BvLz<sub>m</sub>* lines (Table 3), we next investigated the optimal fertilization regime needed to sustain the additional growth and high levels of *BvLz<sub>m</sub>* production. Four representative triple promoter:*BvLz<sub>m</sub>* lines (2-month old) were subjected to two mineral nutrient supply regimes namely, low fertility (LF or 2.4 mg N per plant, twice a week) and a high fertility (HF or 8 mg N per plant), using a balanced commercial fertilizer (Peters Professional 20–20–20; see “Materials and methods” section). *BvLz<sub>m</sub>* yield and growth traits were measured at 2-, 6-, and 8-months following fertilization. Supplemental fertilization increased culm biomass and *BvLz<sub>m</sub>* yield in the triple promoter:*BvLz<sub>m</sub>* lines over time. The most significant increases ( $p < 0.05$ ) between LF and HF were noted at 2 months for all lines (Fig. 5). For instance, pUPE:*BvLz<sub>m</sub>* line 32C (CP72-1210 variety) and pUDE:*BvLz<sub>m</sub>* lines 19, 44 and 54 (TCP98-4454 variety) showed 4.6-, 2.5-, 3.0- and 2.0-fold increases in culm biomass and 1.7-, 1.3-, 1.1 and 1.0-fold enhancements in *BvLz<sub>m</sub>* yield, respectively.

Leaf macronutrient contents of the triple promoter:*BvLz<sub>m</sub>* plants were also monitored following growth under the two fertilization regimes. Plants grown under high nutrient supply rates had significantly ( $p < 0.0001$ ) higher leaf mineral nutrient contents compared to those grown under low nutrient supply rates (Table 4). Leaves of HF plants had higher levels of N, phosphorus (P), potassium (K) and magnesium (Mg), compared to leaves of LF plants (Table 4). In general, leaf nutrient content of the *BvLz<sub>m</sub>* expressing lines was improved by supplemental fertilization, resulting in a 1.5- to 2.2-fold increase in culm biomass and a subsequent 1.2- to 2.2-fold enhancement in *BvLz<sub>m</sub>* yield at 8 month-growth stage (Fig. 5). Taken together, the accumulation of *BvLz<sub>m</sub>* in response to fertilization and the ontogenic *BvLz<sub>m</sub>* accumulation pattern underscore the need for adequate input availability to sustain not only biomass production but also the yield of high-value proteins in crops such as sugarcane.



**Figure 4.** Temporal pattern of recombinant bovine lysozyme (BvLz<sub>m</sub>) accumulation in culms of single promoter:*BvLz<sub>m</sub>* expressing sugarcane lines. BvLz<sub>m</sub> activity of four representative maize *ubiquitin 1* promoter:*BvLz<sub>m</sub>* lines is shown as determined by enzyme-linked immunosorbent assay in 200.0 ml of juice extract from the 7- and 9-month-harvests, and 650.0–700.0 ml of juice extract from the 11-month-harvest (one kg of culm for all harvests). Values represent four biological samples for each *BvLz<sub>m</sub>* expressing line and are reported with standard errors from three technical replications. *BvLz<sub>m</sub>*: maize codon-optimized *BvLz*. Values with different letters are significantly different ( $p < 0.05$ ).

## Conclusions

The genetic/biotechnology tools and resources developed in this study not only expands the utility of sugarcane for large-scale production of recombinant proteins but can be utilized with other monocots and bioenergy feedstocks. Our approach comprises stacking multiple promoters to co-express codon-optimized recombinant genes from different expression vectors using combinatorial transformation methods. This resulted in high recombinant protein yield (up to 11.5% of TSP or 82.5 mg/kg) in transgenic culms, rendering it an attractive biopharming tool for potential commercial uses<sup>69</sup>. We also showed that recombinant BvLz<sub>m</sub> levels can be maintained stably throughout the growing season and had no negative consequences on sugarcane agronomic performance. Overall, our study provides new knowledge, tools and resources to expand the utility of sugarcane beyond a food crop and bioenergy feedstock to using it as a biofactory for expressing high-value proteins<sup>25</sup>.

## Materials and methods

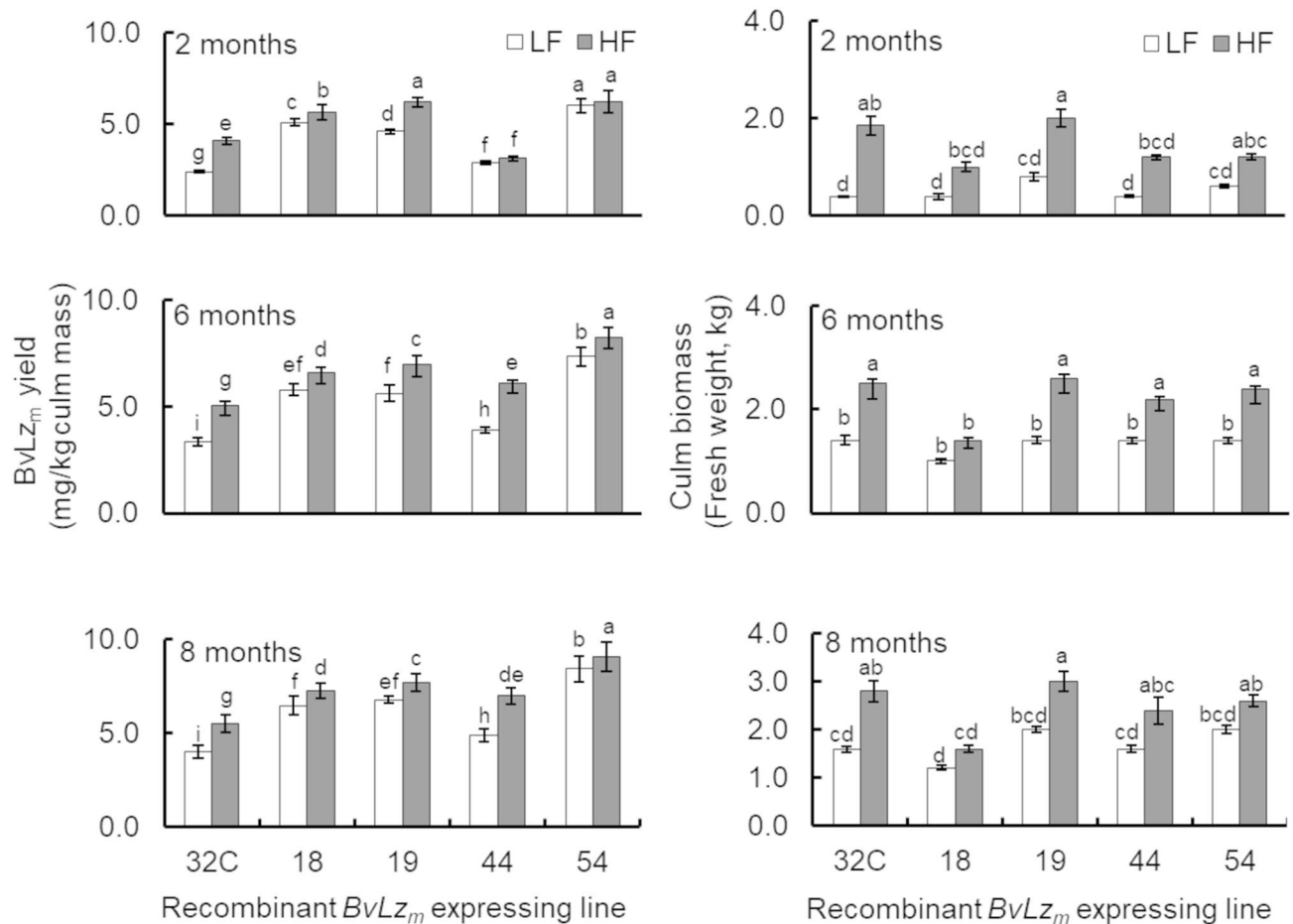
**Expression vectors.** *Basic vectors.* A series of expression vectors were constructed, using a custom synthesized bovine lysozyme (*BvLz*) gene codon-optimized for expression in maize (*BvLz<sub>m</sub>*) (444.0 base pairs [bp])<sup>39</sup> (GenScript, Piscataway, NJ).

The *BvLz<sub>m</sub>* gene was subcloned into pUC57 at *Bam*HI and cloned at the same site into pZero2 (Invitrogen, ThermoFisher Scientific, Waltham, MA), to which the 35ST<sup>34,38,39</sup> (197.0 bp) was added at the *Pst*I site, resulting in the *BvLz<sub>m</sub>*-35ST/pZero2 plasmid.

Three basic *BvLz* expression vectors were generated with the constitutive promoters pUbi<sup>33</sup>, pSHPRP<sup>36</sup> or pSHEF1α<sup>36</sup>. The first vector, pUbi-*BvLz<sub>m</sub>*-35ST/pZero2 was produced by cloning the pUbi fragment (1,977 bp), released from pAHC20 (pUbi:BAR/pUC8)<sup>70</sup> (pUbi minus heat shock element; a 28.0 bp deletion at the 5' end of pUbi) with *Bam*HI/*Hind*III and filled in, into the filled-in *BvLz<sub>m</sub>*-35ST/pZero2. For the other two vectors, the *Sma*I-treated pSHPRP (3,016 bp) and pSHEF1α (1,959 bp) fragments from pSK<sup>+</sup><sup>36</sup> were fused to the *Sna*BI/*Bbs*I-treated/filled-in *BvLz<sub>m</sub>*-35ST fragment from pUbi-*BvLz<sub>m</sub>*-35ST/pZero2 to yield pSHPRP-*BvLz<sub>m</sub>*-35ST/pSK<sup>+</sup> and pSHEF1α-*BvLz<sub>m</sub>*-35ST/pSK<sup>+</sup>, respectively.

Two basic *BvLz* expression vectors were generated with the culm-regulated promoters pSHDIR16<sup>35</sup> or pSCVB21<sup>34</sup>. The pSHDIR16-*BvLz<sub>m</sub>*-35ST/pSK<sup>+</sup> vector was assembled by fusing *BvLz<sub>m</sub>*-35ST, excised from *Bam*HI/*Eco*RI-treated *BvLz<sub>m</sub>*-35ST/pZero2, to the pSHDIR16 fragment<sup>35</sup> (2,680 bp) at the same sites in pSK<sup>+</sup>. The pSCVB21-*BvLz<sub>m</sub>*-35ST/pGEMT-T Easy vector was produced by cloning *BvLz<sub>m</sub>*-35ST, excised from *Bam*HI/*Eco*RI-treated *BvLz<sub>m</sub>*-35ST/pZero2, into the *Nco*I-treated/filled-in pSCVB21 (1,816 bp)/pGEM-T Easy<sup>34</sup>.

*Double terminator vectors.* *BvLz* constructs with a double terminator were generated by fusing the NOST (253 bp)<sup>39</sup> to the 35ST of basic *BvLz* constructs. The pUbi-*BvLz<sub>m</sub>*-35STNOST/pZero2 vector was constructed by releasing the NOST from pBI221 (Accession Number AF502128) (Clontech Laboratories, Inc., Mountain View, CA) with *Eco*RI/*Sst*I, filled in and cloned into the *Xho*I-treated/filled-in pUbi-*BvLz<sub>m</sub>*-35ST/pZero2. To



**Figure 5.** Enhancement of culm biomass and yield of recombinant bovine lysozyme (BvLz<sub>m</sub>) by fertilization in triple promoter: BvLz<sub>m</sub> sugarcane lines. BvLz<sub>m</sub> activity of four representative lines is shown as determined by enzyme-linked immunosorbent assay in juice extract of culms (1.0 kg of culm). Values represented four biological samples and three technical replications at 2, 6 and 8 months following low (LF) or high (HF) fertilization. Values with different letters are significantly different ( $p < 0.05$ ). BvLz<sub>m</sub>, maize codon-optimized BvLz; 32C, pUPE: BvLz<sub>m</sub> line; 18, 44 and 54, pUDE: BvLz<sub>m</sub> lines; U, maize ubiquitin 1 promoter; P, sugarcane proline-rich protein promoter; E, sugarcane elongation factor 1 $\alpha$  promoter; and D, sugarcane dirigent16 promoter.

make pSHPRP-BvLz<sub>m</sub>-35STNOST/pSK<sup>+</sup> and pSHEF1 $\alpha$ -BvLz<sub>m</sub>-35STNOST/pSK, the *Sna*BI/*Bbs*I-treated/filled-in BvLz<sub>m</sub>-35STNOST fragment from pUbi-BvLz<sub>m</sub>-35STNOST/pZero2 was fused to the *Sma*I-treated pSHPRP/pSK<sup>+</sup> and pSHEF1 $\alpha$ /pSK<sup>+</sup> vectors, respectively. To generate the pSCBV21-BvLz<sub>m</sub>-35STNOST/pGEM-T Easy vector, the *Sna*BI/*Bbs*I-treated/filled-in BvLz<sub>m</sub>-35STNOST fragment from pUbi-BvLz<sub>m</sub>-35STNOST/pZero2 was cloned into *Nco*I-treated/filled-in pSCBV21/pGEM-T Easy.

**Vectors with viral untranslated regions.** The 3'UTR of SrMV strain H (GenBank Accession Number U57358) (235.0 bp) was custom synthesized as a fusion to BvLz<sub>m</sub> in pJI (BvLz<sub>m</sub>-SrMV 3'UTR/pJI) (ATUM, DNA2.0, Newark, CA). The pUbi-BvLz<sub>m</sub>-SrMV 3'UTR-35ST/pZero2 vector was assembled by cloning the filled-in SrMV 3'UTR, released from *Eco*RI/*Bgl*II-treated BvLz<sub>m</sub>-3'UTR/pJI, into pUbi-BvLz<sub>m</sub>-35ST/pZero2 at the *Sma*I site. The pSHPRP-BvLz<sub>m</sub>-SrMV 3'UTR-35ST/pSK<sup>+</sup> and pSHEF1 $\alpha$ -BvLz<sub>m</sub>-SrMV 3'UTR-35ST/pSK<sup>+</sup> vectors were generated by fusing the *Sna*BI/*Bbs*I-treated/filled-in BvLz<sub>m</sub>-SrMV 3'UTR-35ST fragment from the pUbi-BvLz<sub>m</sub>-SrMV 3'UTR-35ST/pZero2 to pSHPRP/pSK<sup>+</sup> and pSHEF1 $\alpha$ /pSK<sup>+</sup> at the *Sma*I site, respectively. For construction of pSHDIR16-BvLz<sub>m</sub>-SrMV 3'UTR-35ST/pSK<sup>+</sup> vector, SrMV 3'UTR was released from BvLz<sub>m</sub>-SrMV 3'UTR/pJI by *Eco*RV treatment and cloned into pSHDIR16-BvLz<sub>m</sub>-35ST/pSK<sup>+</sup> at the *Eco*RV site.

All DNA cloning steps were carried out as described by Sambrook<sup>71</sup>. Filling in of endonuclease-treated DNA fragments and dephosphorylation of vectors were done using T4 DNA polymerase (NEB BioLabs, Ipswich, MA) and antarctic phosphatase (NEB BioLabs), respectively.

**Sugarcane transformation.** Tops of field-grown sugarcane (*Saccharum* spp. hybrids) commercial varieties CP72-1210, CP84-1198, TCP87-3388 and TCP98-4454 were collected during the growing season, and leaf roll discs were prepared for stable transformations as previously described<sup>72</sup>. Briefly, leaf blades and sheaths were removed down to the top visible dewlap leaf, and the upper 20–30 cm portion of shoot (leaf roll culm) was

Line/treatment	Macronutrient content (mg/g tissue dry mass)			
	Nitrogen	Phosphorus	Potassium	Magnesium
pUPE: <i>BvLz<sub>m</sub></i> 32C				
LF	8.8±1.0 <sup>a</sup>	1.5±0.2 <sup>a</sup>	11.7±0.2 <sup>a</sup>	1.4±0.4 <sup>a</sup>
HF	15.0±1.0 <sup>b</sup>	3.1±0.4 <sup>b</sup>	19.5±0.2 <sup>b</sup>	3.1±0.3 <sup>b</sup>
pUDE: <i>BvLz<sub>m</sub></i> 18				
LF	10.3±0.1 <sup>a</sup>	2.2±0.2 <sup>a</sup>	17.3±0.3 <sup>a</sup>	1.1±0.1 <sup>a</sup>
HF	18.7±2.0 <sup>b</sup>	4.0±0.2 <sup>b</sup>	20.0±1.8 <sup>b</sup>	2.1±0.2 <sup>b</sup>
pUDE: <i>BvLz<sub>m</sub></i> 19				
LF	12.4±2.0 <sup>a</sup>	2.1±0.2 <sup>a</sup>	17.3±0.5 <sup>a</sup>	1.1±0.1 <sup>a</sup>
HF	16.8±1.0 <sup>b</sup>	3.3±0.3 <sup>b</sup>	19.5±0.7 <sup>b</sup>	3.0±0.2 <sup>b</sup>
Non-transformed				
LF	9.6±1.5 <sup>a</sup>	1.7±0.2 <sup>a</sup>	13.0±0.4 <sup>a</sup>	1.8±0.2 <sup>a</sup>
HF	15.0±2.1 <sup>b</sup>	2.9±0.4 <sup>b</sup>	15.2±1.4 <sup>b</sup>	3.1±0.3 <sup>b</sup>

**Table 4.** Leaf nutrient contents in leaves of triple promoter:*bovine lysozyme* expressing sugarcane lines grown under two fertilization levels. Leaf tissue was sampled from 8-month-old plants of three representative lines. Values represent means from three biological samples ± standard errors. Means are compared column-wise. Values with different letters are significantly different ( $p < 0.0001$ ). *BvLz<sub>m</sub>*: maize codon-optimized *bovine lysozyme* (*BvLz*); pUPE:*BvLz<sub>m</sub>*: *BvLz<sub>m</sub>* expressed from three promoters, maize *ubiquitin 1*, sugarcane *proline-rich protein* and sugarcane *elongation factor 1α* (pSHEF1α); pUDE:*BvLz<sub>m</sub>*: *BvLz<sub>m</sub>* expressed from three promoters, maize *ubiquitin 1*, sugarcane *dirigent16* and pSHEF1α; HF: high fertilization rate; LF: low fertilization rate.

surface sterilized in 70.0% (v/v) ethanol for 20 min. Immature leaf rolls close to the apical meristem were sliced transversely into 1.0 mm thick sections and cultured on MS3 medium (MS medium with 3.0 mg/l of 2,4-dichlorophenoxyacetic acid [2,4-D]) for 30–35 days (for embryogenic calli) or MS0.6 medium (MS with 0.6 mg/l of 2,4-D) for 7–10 days (for embryogenic leaf roll discs). Embryogenic calli and leaf roll discs were preconditioned on MS3- and MS0.6-osmoticum (MS3 or MS0.6 with 0.2 M D-mannitol and 0.2 M D-sorbitol), respectively, for 4 h before and after DNA particle bombardment. DNA bombardment was performed according to Beyene and colleagues<sup>38</sup>. Briefly, tungsten particles (1.1 μm; Bio-Rad Laboratories, Inc.) (1.0 mg) were coated separately with plasmid DNA (1.0 μg) of different constructs at equimolar ratios together with pUbi:BAR/pUC8 selectable marker plasmid using calcium chloride (NaCl) (1.0 M) and spermidine (14.0 mM). The DNA particle suspension (containing the selectable marker plasmid with one or more *BvLz<sub>m</sub>* plasmids) (4.0 μl; 0.5 μg DNA per bombardment) was placed at the center of a syringe filter and delivered into tissue with a particle inflow gun using a 26.0-inch Hg vacuum and a 7.0-cm target distance. Bombarded embryogenic calli and leaf roll discs were maintained on MS3 and MS0.6, respectively, for 10 days in the dark at 28 °C for recovery. They were later incubated in the dark at 28 °C on selection medium (MS3 or MS0.6 with bialaphos at 3.0 mg/l) for a total of 2 weeks. Shoot regeneration and root initiation were performed under bialaphos selection as previously described<sup>72</sup>. Rooted plantlets were transferred to potting soil (Sunshine Mix #1; SunGro Horticulture Distribution, Inc., Agawan, MA) in pots and maintained in the greenhouse.

**Transgenic plant screening.** *Integration and size determination of BvLz<sub>m</sub> expression cassettes.* Integration and size of each *BvLz<sub>m</sub>* expression cassette in the single and multiple stacked promoter:*BvLz<sub>m</sub>* sugarcane lines were determined by Southern blot and PCR analyses, respectively, using genomic DNA isolated according to Tai and Tanksley<sup>73</sup> from liquid N-ground tissues (3.0 g) collected from young leaves of 3–4 month-old plants. Controls included vector-transformed lines and non-transformed plants (tissue culture-derived).

For Southern blot analysis, genomic DNA (10.0 μg per lane) was treated with *Hind*III endonuclease, electrophoresed on 0.8% (w/v) agarose gels and transferred to nylon membranes (Amersham Hybond-XL, GE Healthcare Bio-Sciences Corp., Piscataway, NJ) in 0.4 M sodium hydroxide<sup>74</sup>. Pre-hybridization, hybridization, washing and detection of DNA gel blots were performed using Church's buffer<sup>75</sup>. The probe, corresponding to the *BvLz<sub>m</sub>* coding sequence was amplified by PCR from pUbi-*BvLz<sub>m</sub>*-35ST/pZero2 using the primer set *BvLz*-1F (5'-ATGGCGGCCCTGGTGATCCTGGGCT-3') and *BvLz*-481R (5'-TCACAGGGTGCAGCCTTCCACG-3') and labeled with [ $\alpha$ -<sup>32</sup>P] dCTP using the Random DNA Labeling kit (Invitrogen, ThermoFisher Scientific).

PCR was performed on a C1000 Touch thermal cycler (Bio-Rad Laboratories, Inc., Hercules, CA) in a total reaction volume of 25.0 μl using 200.0 ng of DNA and Platinum *Taq* DNA polymerase (Invitrogen, ThermoFisher Scientific) according to the manufacturer's instructions with the following conditions: 94 °C for 4 min, 35 cycles each at 94 °C for 30 s, 49.7–54.4 °C for 30 s, and 72 °C for 6 min. Primers encompassing the entire promoter:*BvLz<sub>m</sub>*-terminator cassette (Supplementary Table S4) were designed with Primer 3.0. All PCR amplicons were separated by electrophoresis on 0.7% agarose (w/v) gels stained with ethidium bromide. A “no DNA template” was included as a negative control for PCR.

**Determination of *BvLz<sub>m</sub>* copy number.** *BvLz<sub>m</sub>* copy number in single and multiple stacked promoter:*BvLz<sub>m</sub>* sugarcane lines was estimated by qPCR. qPCR was performed on a CFX384 Real-time PCR Detection System (Bio-Rad Laboratories, Inc.) using *iTaq* Universal SYBR Green Supermix (Bio-Rad Laboratories, Inc.), 0.4 μM of each target specific primer and 1.0 ng of genomic DNA from representative transgenic *BvLz<sub>m</sub>* lines, according to the manufacturer's instructions. Primers specific to the promoter-*BvLz<sub>m</sub>* gene junction area (Supplementary Table S4) were designed with Primer 3.0 (<https://bioinfo.ut.ee/primer3-0.4.0/primer3/>). qPCR conditions were as follows: 95.0 °C for 3 min, 39 two-step cycles each at 96.0 °C for 5 s and 57 °C for 30 s, and a final melting curve of 60.0 °C to 95.0 °C for 6 min. The sugarcane anthranilate phosphoribosyltransferase and prolyl 4-hydroxylase genes were used as a reference for single copy genes<sup>76</sup>. qPCR was performed twice in triplicate with two biological replications. PCR efficiency was calculated with LinReg<sup>77</sup>. Results were analyzed and recorded as  $C_T$  (threshold cycle) values. Copy number of the *BvLz<sub>m</sub>* gene was estimated by qPCR according to Casu et al.<sup>76</sup> using the formula  $GCI = \text{Eff}^{\text{ref}C_T} / \text{Eff}^{C_T}$ , where: GCI = gene copy number index,  $\text{Eff}^{\text{ref}C_T}$  = PCR efficiency using the reference gene primers to the power of the reference gene  $C_T$  value for each sample, and  $\text{Eff}^{C_T}$  = PCR efficiency using the test gene primers to the power of the test gene  $C_T$  value generated for each sample.

**Expression analysis of *BvLz<sub>m</sub>*.** Total RNA was isolated by grinding 1.0 g of young leaves collected from 3–4 month-old plants in liquid N<sup>39,78</sup>. For northern blot analysis, RNA (15.0 μg per lane) was fractionated on 1.6% formaldehyde agarose denaturing gels in HEPES buffer and blotted onto nylon membranes (Amersham Hybond-XL) in 10x SSC<sup>75</sup>. Pre-hybridization, *BvLz<sub>m</sub>* probe labeling, hybridization, washing and detection of RNA gel blots were performed as described for Southern blot analysis.

**Plant growth and treatment conditions.** For growth cycle investigations, single-node culm cuttings of 15 single promoter pU:*BvLz<sub>m</sub>* transgenic lines and non-transformed plants were pre-germinated in seedling flats (Supplementary Fig. S1) for 2.5 weeks and transplanted into 37.0-l pots (four pots per line) in commercial growth medium (Sunshine Mix #1). Plants were maintained in a temperature-regulated greenhouse with average day/night temperatures of 32/22 °C and relative humidity of 60–100%. Plants were initially fertilized once per week with a commercial high-phosphorus soluble fertilizer (Peters 8%N-19.8%P-12.5%K; The Scotts Company, Marysville, OH) for 5 weeks and then with a balanced/complete soluble fertilizer (Peters Professional 20–20–20; The Scotts Company) containing N 200.0 g/kg, P 80.0 g/kg, K 166.0 g/kg, Mg 1.0 g/kg, iron 0.5 g/kg, manganese 0.3 g/kg, boron 0.1 g/kg, copper 0.13 g/kg, molybdenum 0.05 g/kg, and zinc 0.25 g/kg.

To assess the impacts of mineral nutrient supply on growth and *BvLz<sub>m</sub>* accumulation, plants from four representative triple promoter pUDE:*BvLz<sub>m</sub>* lines and one representative triple promoter pUPE:*BvLz<sub>m</sub>* line were pre-germinated and transplanted into 15.0-l plastic pots containing the same growth medium as described above. All pots were initially fertilized with a high-phosphorus fertilizer (Peters 8%N-19.8%P-12.5%K; Scotts, Marysville, OH; equivalent to 10.0 kg N/ha). After 2 months, pots were randomly assigned into two fertilization treatment groups, namely, high fertility (HF) and low fertility (LF), with four pots per line selected for each group. Non-transformed plants (tissue culture-derived) were included as negative controls. Fertilization treatments were achieved with a complete fertilizer (Peters Professional 20–20–20) containing macro- and micro-nutrients as described above. Plants in the LF group received an additional equivalent of 20.0 kg N/ha whereas HF plants received 50.0 kg N/ha from supplemental fertilization using Peters Professional 20–20–20 (described above). Fertilizer treatments were applied in split doses (twice per week). Transgenic culms were harvested at 2, 6 and 8 months following fertilization, processed, and their *BvLz* yield was determined by ELISA at the BioSeparation Facility of Texas A&M University's Biological and Agricultural Engineering Department (College Station, Texas).

**Plant physiological analysis.** For inorganic mineral analysis, leaf tissue samples were collected, dried (70 °C for 48 h), ground to pass a 40-μm screen and analyzed for inorganic minerals. Total Kjeldahl N (ammonia and organic N) was determined in digested samples using the EasyChem Plus Analyzer and protocols (Systea Scientific, Chicago, IL), whereas other macronutrients such as P, K and Mg were analyzed using the Optima 7300 DV Inductively Coupled Plasma-Optical Emission Spectrometer (PerkinElmer, Shelton, CT) after partial digestion (hydrolysis) on a HotBlock Digestion System (Environmental Express, Inc., Charleston, SC).

**Total protein extraction.** Large-scale extraction and size fractionation of total soluble proteins (TSPs) from culms (300.0 lbs) of *BvLz<sub>m</sub>* transgenic sugarcane were performed at our Pilot Plant Facility mainly as described previously<sup>79</sup>. Bench-scale extraction and purification of *BvLz<sub>m</sub>* from extracts of transgenic sugarcane culms (100.0 g), using a single-step hydrophobic interaction chromatography, were performed at our BioSeparation Facility (College Station, Texas) as previously described<sup>30</sup>.

For small-scale extraction of TSP from *BvLz<sub>m</sub>* transgenic sugarcane leaf tissue (200.0 mg) was homogenized in 600.0 μl of sodium acetate buffer (50 mM NaOAc, pH 4.4, 0.1 M NaCl) in 2.0 ml tubes for 30 s at 5,000 rpm with the Precellys 24 homogenizer (MO BIO Laboratories, Carlsbad, CA) using ceramic spherical beads (0.64 cm-diameter). TSP supernatants were collected by centrifugation at 13,000g for 25 min at 4 °C.

**Determination of *BvLz<sub>m</sub>* accumulation by enzyme activity and enzyme-linked immunosorbent assays.** To determine the levels of recombinant *BvLz<sub>m</sub>*, enzyme activity and enzyme-linked immunosorbent assays (ELISA) were performed on TSP from culm extract juice. Juice was extracted from 1.0 kg of culms of greenhouse grown *BvLz<sub>m</sub>* transgenic plants at 7, 9 and 11 months for the growth cycle experiment and at 2, 6 and 8 months for the fertilization experiment. For enzyme activity determination, culm extract juice was tested for its ability to lyse *Micrococcus lysodeikticus* cells using the standard protocol from Sigma-Aldrich (St. Louis, MO). Rabbit anti-*BvLz* antibody used in the ELISA was synthesized by Bethyl Laboratories, Inc. (Montgomery,



TX) using tobacco-derived BvLz<sup>31</sup> and further purified through an SP-Sepharose column (GE Healthcare, Piscataway, NJ). ELISA of culm extract juice was performed as previously described<sup>30</sup>. Briefly, a sandwich ELISA consisting of anti-BvLz antibody was used to capture BvLz in juice. Detection was performed using a biotinylated anti-BvLz antibody and horseradish peroxidase-labeled NeutrAvidin (Pierce, ThermoFisher Scientific). The standard curve was generated using BvLz produced in *Pichia pastoris* as in Digan et al.<sup>80</sup>.

**Statistical analysis.** Agronomic data were collected from 3 to 4 independent experiments, with 3–4 replicates per experiment and subjected to an analysis of variance (ANOVA) using the General Linear Model procedure of the Statistical Analysis System 9.4 (SAS Institute Inc., Cary, NC). Mean separation was performed using the Student–Newman–Keuls (SNK) test.

Received: 6 December 2019; Accepted: 21 July 2020

Published online: 13 August 2020

## References

- Demain, A. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* **27**, 297–306. <https://doi.org/10.1016/j.biotechadv.2009.01.008> (2009).
- Basaran, P. & Rodríguez-Cerezo, E. Plant molecular farming: opportunities and challenges. *Crit. Rev. Biotechnol.* **28**, 153–172. <https://doi.org/10.1080/07388550802046624> (2008).
- Rybicki, E. P. Plant-made vaccines for humans and animals. *Plant Biotechnol. J.* **8**, 620–637 (2010).
- Wiktorek-Smagur, A. et al. Green way of biomedicine—how to force plants to produce new important proteins. In *Transgenic Plants—Advances and Limitations* (ed. Çiftçi, Y. O.) 63–90 (InTech, Rijeka, 2012).
- Twyman, R. M., Stoger, E., Schillberg, S., Christou, P. & Fischer, R. Molecular farming in plants: host systems and expression technology. *Trends Biotechnol.* **21**, 570–578 (2003).
- Brumbley, S. M., Purnell, M. P., Petrasovits, L. A., Nielsen, L. K. & Twine, P. H. Developing the sugarcane biofactory for high-value biomaterials. *Int. Sugar J.* **1297**, 5 (2007).
- Ando, S. et al. Overwintering ability and dry matter production of sugarcane hybrids and relatives in the Kanto Region of Japan. *Jpn. Agric. Res. Q.* **45**, 259–267. <https://doi.org/10.6090/jarq.45.259> (2011).
- Matsuoka, S., Kennedy, A. J., Santos, E. G. D., Tomazela, A. L. & Rubio, L. C. S. Energy cane: its concept, development, characteristics, and prospects. *Adv. Bot.* **2014**, 597275 (2014).
- Gallo-Meagher, M. & Irvine, J. E. Effects of tissue type and promoter strength on transient GUS expression in sugarcane following particle bombardment. *Plant Cell Rep.* **12**, 666–670. <https://doi.org/10.1007/BF00233416> (1993).
- Lakshmanan, P. et al. Sugarcane biotechnology: the challenges and opportunities. *Vitro Cell. Dev. Biol. Plant* **41**, 345–363 (2005).
- Anderson, D. J. et al. Synthesis of short-chain-length/medium-chain length polyhydroxyalkanoate (PHA) copolymers in peroxisomes of transgenic sugarcane plants. *Trop. Plant Biol.* **4**, 170–184. <https://doi.org/10.1007/s12042-011-9080-7> (2011).
- McQualter, R. B. et al. Initial evaluation of sugarcane as a production platform for p-hydroxybenzoic acid. *Plant Biotechnol. J.* **3**, 29–41 (2005).
- Petrasovits, L. A. et al. Enhanced polyhydroxybutyrate production in transgenic sugarcane. *Plant Biotechnol. J.* **10**, 569–578 (2012).
- Purnell, M. P., Petrasovits, L. A., Nielsen, L. K. & Brumbley, S. M. Spatio-temporal characterization of polyhydroxybutyrate accumulation in sugarcane. *Plant Biotechnol. J.* **5**, 173–184 (2007).
- Tilbrook, K., Gebbie, L., Schenk, P. M., Poirier, Y. & Brumbley, S. M. Peroxisomal polyhydroxyalkanoate biosynthesis is a promising strategy for bioplastic production in high biomass crops. *Plant Biotechnol. J.* **9**, 958–969 (2011).
- Chong, B. F. et al. Co-ordinated synthesis of gentiobitol and sorbitol, evidence of sorbitol glycosylation in transgenic sugarcane. *Phytochemistry* **71**, 736–741. <https://doi.org/10.1016/j.phytochem.2010.01.014> (2010).
- Mudge, S. R. et al. Mature-stem expression of a silencing-resistant sucrose isomerase gene drives isomaltulose accumulation to high levels in sugarcane. *Plant Biotechnol. J.* **11**, 502–509 (2013).
- Wu, L. & Birch, R. G. Doubled sugar content in sugarcane plants modified to produce a sucrose isomer. *Plant Biotechnol. J.* **5**, 109–117. <https://doi.org/10.1111/j.1467-7652.2006.00224.x> (2007).
- Wang, M.-L., Goldstein, C., Su, W., Moore, P. H. & Albert, H. H. Production of biologically active GM-CSF in sugarcane: a secure biofactory. *Transgenic Res.* **14**, 167–178 (2005).
- Henrique-Silva, F. & Soares-Costa, A. *Transgenic Plants* 437–450 (Springer, Berlin, 2012).
- Gianotti, A. et al. Recombinant expression, purification, and functional analysis of two novel cystatins from sugarcane (*Saccharum officinarum*). *Protein Expr. Purif.* **47**, 483–489. <https://doi.org/10.1016/j.pep.2005.10.026> (2006).
- Ribeiro, C. W. et al. Production of a His-tagged canecystatin in transgenic sugarcane and subsequent purification. *Biotechnol. Prog.* **24**, 1060–1066 (2008).
- Harrison, M. D. et al. Accumulation of recombinant cellobiohydrolase and endoglucanase in the leaves of mature transgenic sugar cane. *Plant Biotechnol. J.* **9**, 884–896. <https://doi.org/10.1111/j.1467-7652.2011.00597.x> (2011).
- Harrison, M. D. et al. Recombinant cellulase accumulation in the leaves of mature, vegetatively propagated transgenic sugarcane. *Mol. Biotechnol.* **56**, 795–802 (2014).
- Palaniswamy, H. et al. Vacuolar targeting of r-proteins in sugarcane leads to higher levels of purifiable commercially equivalent recombinant proteins in cane juice. *Plant Biotechnol. J.* **14**, 791–807 (2016).
- Jolles, P. et al. Stomach lysozymes of ruminants. II. Amino acid sequence of cow lysozyme 2 and immunological comparisons with other lysozymes. *J. Biol. Chem.* **259**, 11617–11625 (1984).
- Sahoo, N. R. et al. Lysozyme in livestock: a guide to selection for disease resistance: a review. *J. Anim. Sci. Adv.* **2**, 347–360 (2012).
- Lemos, F. J. A., Ribeiro, A. F. & Terra, W. R. A bacteria-digesting midgut-lysozyme from *Musca domestica* (Diptera) larvae. Purification, properties and secretory mechanism. *Insect Biochem. Mol. Biol.* **23**, 533–541 (1993).
- Mirkov, T. E. & Fitzmaurice, L. C. Google patents (1995).
- Barros, G. O. F. et al. Recovery of bovine lysozyme from transgenic sugarcane stalks: extraction, membrane filtration, and purification. *Bioprocess Biosyst. Eng.* **36**, 1407–1416. <https://doi.org/10.1007/s00449-012-0878-y> (2013).
- Wilcox, C. P. et al. Production and purification of an active bovine lysozyme in tobacco (*Nicotiana tabacum*): utilization of value-added crop plants traditionally grown under intensive agriculture. *J. Agric. Food Chem.* **45**, 2793–2797 (1997).
- Bock, R. Strategies for metabolic pathway engineering with multiple transgenes. *Plant Mol. Biol.* **83**, 21–31. <https://doi.org/10.1007/s11103-013-0045-0> (2013).

33. Christensen, A. H., Sharrock, R. A. & Quail, P. H. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol. Biol.* **18**, 675–689. <https://doi.org/10.1007/BF0020010> (1992).
34. Gao, S.-J. *et al.* A novel Sugarcane bacilliform virus promoter confers gene expression preferentially in the vascular bundle and storage parenchyma of the sugarcane culm. *Biotechnol. Biofuels* **10**, 172–172. <https://doi.org/10.1186/s13068-017-0850-9> (2017).
35. Damaj, M. B. *et al.* Sugarcane DIRIGENT and O-METHYLTRANSFERASE promoters confer stem-regulated gene expression in diverse monocots. *Planta* **231**, 1439–1458. <https://doi.org/10.1007/s00425-010-1138-5> (2010).
36. Yang, M., Bower, R., Burow, M. D., Paterson, A. H. & Mirkov, T. E. A rapid and direct approach to identify promoters that confer high levels of gene expression in monocots. *Crop Sci.* **43**, 1805–1813 (2003).
37. Woodard, S. L. *et al.* American Society of Agricultural and Biological Engineers.
38. Beyene, G. *et al.* Unprecedented enhancement of transient gene expression from minimal cassettes using a double terminator. *Plant Cell Rep.* **30**, 13–25. <https://doi.org/10.1007/s00299-010-0936-3> (2011).
39. Damaj, M. B. & Mirkov, T. E. Compositions, organisms, systems, and methods for expressing a gene product in plants. U.S. patent (2019).
40. Moon, K.-B. *et al.* Development of systems for the production of plant-derived biopharmaceuticals. *Plants* **9**, 30 (2020).
41. Merlin, M., Gecchele, E., Capaldi, S., Pezzotti, M. & Avesani, L. Comparative evaluation of recombinant protein production in different biofactories: the green perspective. *BioMed Res. Int.* **2014**, 1–14 (2014).
42. Jiang, M.-C., Hu, C.-C., Lin, N.-S. & Hsu, Y.-H. Production of human IFN $\gamma$  protein in *Nicotiana benthamiana* plant through an enhanced expression system based on bamboo mosaic virus. *Viruses* **11**, 509 (2019).
43. Zischewski, J., Sack, M. & Fischer, R. Overcoming low yields of plant-made antibodies by a protein engineering approach. *Biotechnol. J.* **11**, 107–116 (2016).
44. Buyel, J. F., Twyman, R. M. & Fischer, R. Very-large-scale production of antibodies in plants: the biologization of manufacturing. *Biotechnol. Adv.* **35**, 458–465 (2017).
45. Sainsbury, F. *et al.* Rapid transient production in plants by replicating and non-replicating vectors yields high quality functional anti-HIV antibody. *PLoS ONE* **5**, e13976 (2010).
46. Lai, H., He, J., Engle, M., Diamond, M. S. & Chen, Q. Robust production of virus-like particles and monoclonal antibodies with geminiviral replicon vectors in lettuce. *Plant Biotechnol. J.* **10**, 95–104 (2012).
47. Porceddu, A. *et al.* Transgenic plants expressing human glutamic acid decarboxylase (GAD65), a major autoantigen in insulin-dependent diabetes mellitus. *Mol. Breed.* **5**, 553–560 (1999).
48. Uvarova, E. A. *et al.* Oral immunogenicity of plant-made *Mycobacterium tuberculosis* ESAT6 and CFP10. *BioMed Res. Int.* **2013**, 1–13 (2013).
49. Mason, H. S. *et al.* Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice. *Proc. Natl. Acad. Sci.* **93**, 5335–5340 (1996).
50. Zhang, X., Buehner, N. A., Hutson, A. M., Estes, M. K. & Mason, H. S. Tomato is a highly effective vehicle for expression and oral immunization with Norwalk virus capsid protein. *Plant Biotechnol. J.* **4**, 419–432 (2006).
51. Tokuhara, D. *et al.* Rice-based oral antibody fragment prophylaxis and therapy against rotavirus infection. *J. Clin. Investig.* **123**, 3829–3838 (2013).
52. Borghi, L. *Plant Developmental Biology* 65–75 (Springer, Berlin, 2010).
53. Maruyama, K. *et al.* Design of an optimal promoter involved in the heat-induced transcriptional pathway in Arabidopsis, soybean, rice and maize. *Plant J.* **89**, 671–680 (2017).
54. Gatz, C. Chemically inducible promoters in transgenic plants. *Curr. Opin. Biotechnol.* **7**, 168–172 (1996).
55. Butaye, K. M. J., Cammue, B. P. A., Delauré, S. L. & De Bolle, M. F. C. Approaches to minimize variation of transgene expression in plants. *Mol. Breed.* **16**, 79–91. <https://doi.org/10.1007/s11032-005-4929-9> (2005).
56. Kohli, A. *et al.* The quest to understand the basis and mechanisms that control expression of introduced transgenes in crop plants. *Plant Signal. Behav.* **1**, 185–195. <https://doi.org/10.4161/psb.1.4.3195> (2006).
57. Rajeev Kumar, S., Anunanthini, P. & Ramalingam, S. Epigenetic silencing in transgenic plants. *Front. Plant Sci.* **6**, 693. <https://doi.org/10.3389/fpls.2015.00693> (2015).
58. Greger, I. H., Proudfoot, N. J., Demarchi, F. & Giacca, M. Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. *Nucleic Acids Res.* **26**, 1294–1300. <https://doi.org/10.1093/nar/26.5.1294> (1998).
59. Ingelbrecht, I. *et al.* Transcriptional interference in transgenic plants. *Gene* **109**, 239–242 (1991).
60. Kadesch, T. & Berg, P. Effects of the position of the simian virus 40 enhancer on expression of multiple transcription units in a single plasmid. *Mol. Cell. Biol.* **6**, 2593–2601 (1986).
61. Maqbool, S. B. & Christou, P. Multiple traits of agronomic importance in transgenic indica rice plants: analysis of transgene integration patterns, expression levels and stability. *Mol. Breed.* **5**, 471–480 (1999).
62. Shearwin, K. E., Callen, B. P. & Egan, J. B. Transcriptional interference—a crash course. *Trends Genet.* **21**, 339–345 (2005).
63. Horlick, R. A. *et al.* Combinatorial gene expression using multiple episomal vectors. *Gene* **243**, 187–194 (2000).
64. Egelkrout, E. *et al.* Enhanced expression levels of cellulase enzymes using multiple transcription units. *BioEnergy Res.* **6**, 699–710 (2013).
65. Urreta, I. & Castañón, S. Transgenic plants as biofactories for the production of biopharmaceuticals: A case study of human placental lactogen. In *Transgenic Plants-Advances and Limitations* (ed. Çiftçi, Y. O.) 305–328 (InTech, Rijeka, 2012).
66. Legendre, B. L. Ripening of sugarcane: effects of sunlight, temperature, and rainfall I. *Crop Sci.* **15**, 349–352 (1975).
67. Duwadi, K., Chen, L., Menassa, R. & Dhaubhadel, S. Identification, characterization and down-regulation of cysteine protease genes in tobacco for use in recombinant protein production. *PLoS ONE* **10**, e0130556 (2015).
68. Marschner, H. *Mineral Nutrition of Higher Plants* (Academic Press, NY, 1995).
69. Davies, H. M. Review article: commercialization of whole-plant systems for biomanufacturing of protein products: evolution and prospects. *Plant Biotechnol. J.* **8**, 845–861. <https://doi.org/10.1111/j.1467-7652.2010.00550.x> (2010).
70. Christensen, A. H. & Quail, P. H. Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res.* **5**, 213–218. <https://doi.org/10.1007/BF01969712> (1996).
71. Sambrook, H. C. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989).
72. Ramasamy, M. *et al.* A biolistic-based genetic transformation system applicable to a broad-range of sugarcane and energycane varieties. *GM Crops Food* **9**, 211–227 (2018).
73. Tai, T. H. & Tanksley, S. D. A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. *Plant Mol. Biol. Report.* **8**, 297–303 (1990).
74. Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001).
75. Mangwende, T. *et al.* The P0 gene of Sugarcane yellow leaf virus encodes an RNA silencing suppressor with unique activities. *Virology* **384**, 38–50 (2009).
76. Casu, R. E., Selivanova, A. & Perroux, J. M. High-throughput assessment of transgene copy number in sugarcane using real-time quantitative PCR. *Plant Cell Rep.* **31**, 167–177. <https://doi.org/10.1007/s00299-011-1150-7> (2012).
77. Ruijter, J. M. *et al.* Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* **37**, e45–e45 (2009).

78. Damaj, M. B. *et al.* Reproducible RNA preparation from sugarcane and citrus for functional genomic applications. *Int. J. Plant Genom.* **2009**, 13. <https://doi.org/10.1155/2009/765367> (2009).
79. Mirkov, T. E., Monclin, J. P., Barrilleaux, A., Irvine, J. E. & Moonan, F. Google patents (2002).
80. Digan, M. E. *et al.* Continuous production of a novel lysozyme via secretion from the yeast, *Pichia pastoris*. *Biotechnology* **7**, 160 (1989).

## Acknowledgements

We would like to dedicate this manuscript to late Professor T. Erik Mirkov (1959–2018), a pioneer in sugarcane biotechnology. We gratefully acknowledge Xavier Gonzales for the assembly of three genetic constructs, Renesh Bedre for the statistical analysis of data and Sonia Irigoyen (Texas A&M AgriLife Research) for critical review of the manuscript. We are grateful to Denise Rossi, Hyun Park Kang, Ninfa Ramos, Soledad Al-Varez, Gerleene Acuna, Adan Solis and Abigail Cruz for excellent technical assistance. This research was supported by funds from BioCane, Inc., a subsidiary of Grower Research Group LLC (Soledad, CA) and Texas A&M AgriLife Research Grants (124738-96210; 124190-96210) to K.K.M.

## Author contributions

M.B.D., J.L.J., S.L.W., C.V.-B., G.O.F.B., J.M., Z.L.N., and K.K.M. designed the experiments. B.B.D. designed the combinatorial gene expression system and developed total protein enrichment protocols. M.B.D. made the genetic constructs and prepared the manuscript. M.B.D. and J.M. conducted transformation experiments. M.B.D., J.M. and C.V.-B. conducted transgenic plant screening analyses. J.L.J. conducted growth cycle experiments. M.B.D. and J.M. conducted fertilization experiments. S.L.W., G.O.F.B. and S.G.W. conducted protein extraction, ELISA and purification experiments. S.L.W., J.L.J., Z.L.N., and K.K.M. supervised the study and reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-70530-z>.

**Correspondence** and requests for materials should be addressed to M.B.D. or K.K.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020