



Stable Protein Sialylation in *Physcomitrella*

OPEN ACCESS

Edited by:

Els J. M. Van Damme,
Ghent University, Belgium

Reviewed by:

Muriel Bardor,
Université de Rouen, France
Hiroyuki Kajjura,
Osaka University, Japan

*Correspondence:

Eva L. Decker
eva.decker@biologie.uni-freiburg.de

†ORCID:

Lennard L. Bohlender
orcid.org/0000-0002-4599-7807

Juliana Parsons
orcid.org/0000-0001-6261-2342

Sebastian N. W. Hoernstein
orcid.org/0000-0002-2095-689X

Christine Rempfer
orcid.org/0000-0002-0672-3897

Natalia Ruiz-Molina
orcid.org/0000-0002-0170-6000

Friedrich Altmann
orcid.org/0000-0002-0112-7877

Ralf Reski
orcid.org/0000-0002-5496-6711

Eva L. Decker
orcid.org/0000-0002-9151-1361

Specialty section:

This article was submitted to
Plant Biotechnology,
a section of the journal
Frontiers in Plant Science

Received: 24 September 2020

Accepted: 30 November 2020

Published: 18 December 2020

Citation:

Bohlender LL, Parsons J,
Hoernstein SNW, Rempfer C,
Ruiz-Molina N, Lorenz T, Rodríguez
Jahnke F, Figl R, Fode B, Altmann F,
Reski R and Decker EL (2020) Stable
Protein Sialylation in *Physcomitrella*.
Front. Plant Sci. 11:610032.
doi: 10.3389/fpls.2020.610032

Lennard L. Bohlender^{1†}, Juliana Parsons^{1†}, Sebastian N. W. Hoernstein^{1†},
Christine Rempfer^{1,2†}, Natalia Ruiz-Molina^{1†}, Timo Lorenz¹,
Fernando Rodríguez Jahnke^{1,2}, Rudolf Figl³, Benjamin Fode⁴, Friedrich Altmann^{3†},
Ralf Reski^{1,2,5,6†} and Eva L. Decker^{1*†}

¹ Plant Biotechnology, Faculty of Biology, University of Freiburg, Freiburg, Germany, ² Spemann Graduate School of Biology and Medicine, University of Freiburg, Freiburg, Germany, ³ Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria, ⁴ Eleva GmbH, Freiburg, Germany, ⁵ Signalling Research Centres BIOSS and CIBSS, University of Freiburg, Freiburg, Germany, ⁶ Cluster of Excellence livMatS, Freiburg Center for Interactive Materials and Bioinspired Technologies, University of Freiburg, Freiburg, Germany

Recombinantly produced proteins are indispensable tools for medical applications. Since the majority of them are glycoproteins, their *N*-glycosylation profiles are major determinants for their activity, structural properties and safety. For therapeutical applications, a glycosylation pattern adapted to product and treatment requirements is advantageous. *Physcomitrium patens* (*Physcomitrella*, moss) is able to perform highly homogeneous complex-type *N*-glycosylation. Additionally, it has been glyco-engineered to eliminate plant-specific sugar residues by knock-out of the β 1,2-xylosyltransferase and α 1,3-fucosyltransferase genes (Δ xt/ft). Furthermore, *Physcomitrella* meets wide-ranging biopharmaceutical requirements such as GMP compliance, product safety, scalability and outstanding possibilities for precise genome engineering. However, all plants, in contrast to mammals, lack the capability to perform *N*-glycan sialylation. Since sialic acids are a common terminal modification on human *N*-glycans, the property to perform *N*-glycan sialylation is highly desired within the plant-based biopharmaceutical sector. In this study, we present the successful achievement of protein *N*-glycan sialylation in stably transformed *Physcomitrella*. The sialylation ability was achieved in a Δ xt/ft moss line by stable expression of seven mammalian coding sequences combined with targeted organelle-specific localization of the encoded enzymes responsible for the generation of β 1,4-galactosylated acceptor *N*-glycans as well as the synthesis, activation, transport and transfer of sialic acid. Production of free (Neu5Ac) and activated (CMP-Neu5Ac) sialic acid was proven. The glycosidic anchor for the attachment of terminal sialic acid was generated by the introduction of a chimeric human β 1,4-galactosyltransferase gene under the simultaneous knock-out of the gene encoding the endogenous β 1,3-galactosyltransferase. Functional complex-type *N*-glycan sialylation was confirmed via mass spectrometric analysis of a stably co-expressed recombinant human protein.

Keywords: glyco-optimization, *N*-glycan sialylation, *N*-glycan humanization, glyco-engineering, plant-made recombinant biopharmaceuticals, plant-made pharmaceuticals, PMP

INTRODUCTION

The biopharmaceutical sector is continuously increasing its significant share of the global pharmaceutical market (Moorkens et al., 2017). Recombinant proteins, which represent the largest proportion of biopharmaceutical products, are manufactured only by a limited number of production platforms. These are mainly based on bacteria, yeast, insect cell, plant cell or mammalian cell cultures, whereas the latter are the most represented ones (Merlin et al., 2014). Biopharmaceuticals produced in mammalian cell cultures are in general well accepted, since these are able to perform post-translational modifications in a similar manner as in the human body (Jenkins et al., 2008).

Within mammalian systems, Chinese hamster ovary (CHO) cells are by far the most common production host (Walsh, 2018; Nguyen et al., 2020); even though they are naturally not able to link sialic acid to *N*-glycans in the human-typical predominant α 2,6-manner but perform α 2,3-linkages instead. This linkage type is described to be destabilizing on recombinantly produced IgG1 due to steric effects that weaken the glycan–backbone interaction, resulting in a less stable conformation (Zhang et al., 2019). Furthermore, *N*-glycans of glycoproteins produced in some non-human mammalian cell lines can carry immunogenic residues, such as the terminal *N*-glycolylneuraminic acid (Neu5Gc), a sialic acid not produced in humans (Chou et al., 2002; Tangvoranuntakul et al., 2003; Ghaderi et al., 2012) or α 1,3-linked galactoses, which especially occur in SP2/0 cells (Hamadeh et al., 1992; Chung et al., 2008). Moreover, mammalian cell lines show a high heterogeneity in terms of *N*-glycosylation as well as a limited batch-to-batch reproducibility (Jefferis, 2005; Kolarich et al., 2012; Komatsu et al., 2016), which can negatively affect both the efficacy as well as the quality of a biopharmaceutical (Schiestl et al., 2011). Alternative production platforms can offer advantages in sectors neglected by the current narrowing range of systems. In this regard, plants combine the protein processing capabilities of eukaryotic cells with cultivation requirements comparable to prokaryotic production systems, leading to lower manufacturing costs (Rozov et al., 2018). Further, plants lack human pathogens, endotoxins and oncogenic DNA sequences (Commandeur et al., 2003), and hence are generally safer than microbial or animal production hosts (Twyman et al., 2003). Between plants and animals, both the protein biosynthesis as well as the secretory pathway are highly conserved. This enables plants, in contrast to bacterial hosts, to synthesize and fold complex human proteins correctly (Tschofen et al., 2016), as well as to perform the majority of posttranslational modifications needed for high-quality biopharmaceuticals.

Due to their smaller glycome, plants display a reduced *N*-glycosylation microheterogeneity in comparison to mammalian cells (Bosch et al., 2013; Margolin et al., 2020). The latter usually synthesize a wide mixture of *N*-glycans, whereas in plants often one *N*-glycan structure predominates resulting in a generally high homogeneity of *N*-glycosylation (Koprivova et al., 2003; Castilho et al., 2011). This leads to a rigid batch-to-batch stability and an improvement of quality and kinetics of recombinantly plant-produced biopharmaceuticals (Bosch et al., 2013; Stoger et al., 2014; Sochaj et al., 2015; Shen

et al., 2016). Like mammals, plants produce complex-type *N*-glycans, which share an identical $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ (GnGn) core but differ in some post-ER processed residues from their mammalian counterparts (Viëtör et al., 2003). The distal *N*-glycans harbor β 1,4-linked galactoses, which are often capped by a α 2,6-linked sialic acid residue (Varki et al., 2017). In contrast, plant *N*-glycans are predominantly terminated with exposed distal *N*-glycans, which additionally can be decorated with an β 1,3-linked galactose and an α 1,4-linked fucose (Lerouge et al., 1998). This terminal structure [Gal β (1,3)(Fuc α (1,4))GlcNAc], known as Lewis A (Le^a) epitope, appears predominantly on extracellular membrane-bound or soluble glycoproteins, suggesting that it plays a role in cell-to-cell recognition and interaction with pathogens. However, it can also be found to some extent on glycoproteins within the Golgi apparatus (Fitchette et al., 1999; Strasser et al., 2007). The Le^a epitope is, however, a tumor-associated carbohydrate structure in humans (Zhang et al., 2018) and it has been associated with antibody formation (Fitchette et al., 1999; Wilson et al., 2001). Additionally, the Asn-linked *N*-glycan is decorated with an α 1,3-attached fucose in plants, whereas in mammals this residue is α 1,6-linked. Furthermore, in plants the proximal mannose harbors a β 1,2-linked xylose, a sugar which is not present in mammals. Considering that the majority of biopharmaceuticals are glycoproteins and that their *N*-glycosylation plays an important role in their efficacy (Lingg et al., 2012), attention should be paid to *N*-glycosylation quality. Therefore, several model plants have been glyco-engineered in the last decades to obtain a humanized *N*-glycosylation pattern devoid of these structures (van Ree et al., 2000; Bardor et al., 2003; Gomord et al., 2005; Decker and Reski, 2012; Decker et al., 2014). Plant-specific β 1,2-xylosylation and α 1,3-fucosylation of *N*-glycans were eliminated in *Physcomitrella* by targeted knock-outs of the responsible glycosyltransferase (XT and FT) coding genes via homologous recombination (Koprivova et al., 2004). The same genes were knocked down by RNAi in *Lemna minor* (Cox et al., 2006), *Nicotiana benthamiana* (Strasser et al., 2008), *Medicago sativa* (Sourrouille et al., 2008) and *Oryza sativa* (Shin et al., 2011) or knocked out by T-DNA insertion in *Arabidopsis thaliana* (Strasser et al., 2004). More recently, these genes were knocked out in *Nicotiana benthamiana* (Jansing et al., 2018) and *Nicotiana tabacum* BY-2 suspension cells (Hanania et al., 2017; Mercx et al., 2017) by targeting two XT and either four (*N. benthamiana*) or up to five (*N. tabacum*) FT-encoding genes via CRISPR/Cas9 genome editing. Furthermore, the Le^a epitope, which is a rare terminal modification of *Physcomitrella* *N*-glycans, was abolished in this organism by the single knock-out of the gene encoding the responsible β 1,3-galactosyltransferase 1 (GalT3, Pp3c22_470V3.1; Parsons et al., 2012). As proven via mass spectrometry, recombinant erythropoietin (rEPO) produced in *Physcomitrella* with the triple knock-out (KO) of XT, FT, and GalT3 displayed an outstanding homogeneity in the *N*-glycans, with a sharply predominant GnGn pattern (Parsons et al., 2012), providing a suitable platform for further glyco-optimization approaches. These studies not only achieved the elimination of the undesired sugar structures, they also revealed a sufficient plasticity of plants toward glyco-engineering approaches without

observable phenotypic impairments under different cultivation conditions suitable for biopharmaceutical production (reviewed in Montero-Morales and Steinkellner, 2018).

Plant systems are already being used for the production of biopharmaceuticals. β -Glucocerebrosidase, an enzyme for replacement therapy in Morbus Gaucher treatment, is obtained from carrot-cell suspensions (van Dussen et al., 2013) or ZMapp, a combination of antibodies for treatment of Ebola infections, is produced in *N. benthamiana* lacking plant-typical N-glycan modifications (Margolin et al., 2018). Besides these approved plant-made biopharmaceuticals there are further promising candidates in clinical trials, such as hemagglutinin-based virus-like particles as *N. benthamiana*-derived vaccine against influenza (Smith et al., 2020) or the moss-produced α -galactosidase for enzyme replacement therapy in Morbus Fabry treatment (Shen et al., 2016; Hennermann et al., 2019). The majority of pharmaceutically interesting glycoproteins are terminally sialylated in their native form. N-glycan sialylation is highly desirable due to its role in half-life, solubility, stability and receptor binding (Varki et al., 2017). However, plants are unable to perform β 1,4-galactosylation, which in mammals serves as acceptor substrate for terminal N-glycan sialylation. Moreover, they are not able to produce, activate and link sialic acid (Zeleny et al., 2006; Bakker et al., 2008; Castilho et al., 2008), which in mammals requires the coordinated activity of six enzymes. To achieve *in planta* β 1,4-galactosylation, several trials have been undertaken in different systems with varying degrees of success, including the expression of the β 1,4-galactosyltransferase from humans (GalT4) or animals or chimeric varieties (Palacpac et al., 1999; Bakker et al., 2001, 2006; Misaki et al., 2003; Huether et al., 2005; Fujiyama et al., 2007; Hesselink et al., 2014; Kittur et al., 2020; Kriechbaum et al., 2020). The efficiency and quality of galactosylation was shown to depend on the localization of the enzyme in the Golgi apparatus (Bakker et al., 2001; Strasser et al., 2009), on its expression level (Kallolimath et al., 2018) and on the investigated protein (Kriechbaum et al., 2020). *De novo* synthesis of sialic acid involves three enzymes in a four-step process within the cytosol. The biosynthesis starts with the synthesis of N-acetylmannosamine (ManNAc) out of its UDP-activated precursor substrate N-acetylglucosamine (UDP-GlcNAc), which is present in both, plants and mammals. The first reaction steps of the sialic acid production are catalyzed by the enzyme UDP-GlcNAc 2-epimerase/ManNAc kinase (GNE), which is the key enzyme of sialic acid biosynthesis (Reinke et al., 2009). GNE bifunctionally accomplishes the cleavage of UDP and the subsequent epimerization of GlcNAc to ManNAc, followed by ManNAc phosphorylation at C-6, resulting in ManNAc-6P. Subsequently, Neu5Ac-9-phosphate synthase (NANS) catalyzes the condensation with phosphoenolpyruvate (PEP) leading to Neu5Ac-9-P (Roseman et al., 1961). Finally, Neu5Ac-9-P is dephosphorylated by the Neu5Ac-9-P phosphatase (NANP) resulting in N-acetylneuraminic acid (Neu5Ac, sialic acid, Maliyekal et al., 2006). Activation of Neu5Ac is then performed by the nuclear CMP-Neu5Ac synthetase (CMAS), forming CMP-Neu5Ac (activated sialic acid). This activated sialic acid is translocated in exchange for CMP in an antiporter mechanism by the CMP-sialic acid transporter (CSAT) from the cytosol

into the Golgi lumen (Aoki et al., 2003; Tiralongo et al., 2006; Zhao et al., 2006). In the final step of mammalian N-glycosylation, the sialyltransferase (ST) catalyzes the release of Neu5Ac from CMP and the covalent attachment of it to an β 1,4-galactosylated acceptor in the *trans*-Golgi apparatus, resulting in the complex-type mammalian N-glycosylation (pathway depicted in **Supplementary Figure 1A**). In recent years several approaches have been performed to establish functional sialylation in plants. Accordingly, the simultaneous expression of the murine GNE and human NANS and CMAS coding sequences (CDS) in *A. thaliana* led to the production of activated sialic acid (Castilho et al., 2008). In *N. benthamiana* functional N-glycan sialylation could be achieved in a transient approach by co-expressing five CDSs of the sialylation pathway-genes, together with a chimeric GalT4 containing N-terminally the sub-Golgi apparatus localization determining cytoplasmic, transmembrane, and stem (CTS) region of the rat ST. This resulted in a sialylation efficiency of about 80%, detected on an additionally transiently co-expressed recombinant monoclonal antibody (Castilho et al., 2010). The first stable genetically engineered sialylating *N. benthamiana* line was described by Kallolimath et al. (2016).

Particularly, with regard to recent achievements in the production of candidate biopharmaceuticals, combined with its Good Manufacturing Practice (GMP)-compliant production capabilities, the moss *Physcomitrella* serves as a competitive production platform for biopharmaceuticals (Decker and Reski, 2020). Preclinical trials of moss-derived recombinant human complement factor H were recently effectively accomplished (Häffner et al., 2017; Michelfelder et al., 2017), along with the moss-GAA (acid α -1,4-glucosidase) against Pompe disease (Hintze et al., 2020). Furthermore, the clinical phase I study of the recombinantly produced candidate biopharmaceutical moss-aGal against Fabry disease was successfully completed (Reski et al., 2018; Hennermann et al., 2019). Moss-GAA and moss-aGal proved to have better overall performance compared to their variants produced in mammalian cell cultures (Hennermann et al., 2019; Hintze et al., 2020), demonstrating the potential of *Physcomitrella* to produce biobetters. However, the last step in humanizing N-glycans, protein sialylation, still has to be accomplished.

Here, we report the successful establishment of stable protein sialylation in *Physcomitrella*, after subsequently accomplishing the synthesis of free as well as activated sialic acid, and the β 1,4 linkage of the sialic acid anchor galactose to protein N-glycans. This will further increase the attractiveness of moss as plant-based biopharmaceutical production platform.

MATERIALS AND METHODS

Construct Generation

For cloning, coding sequences were amplified by PCR with PhusionTM High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, United States). All primers used in this study are compiled in **Supplementary Table 1**. After each PCR

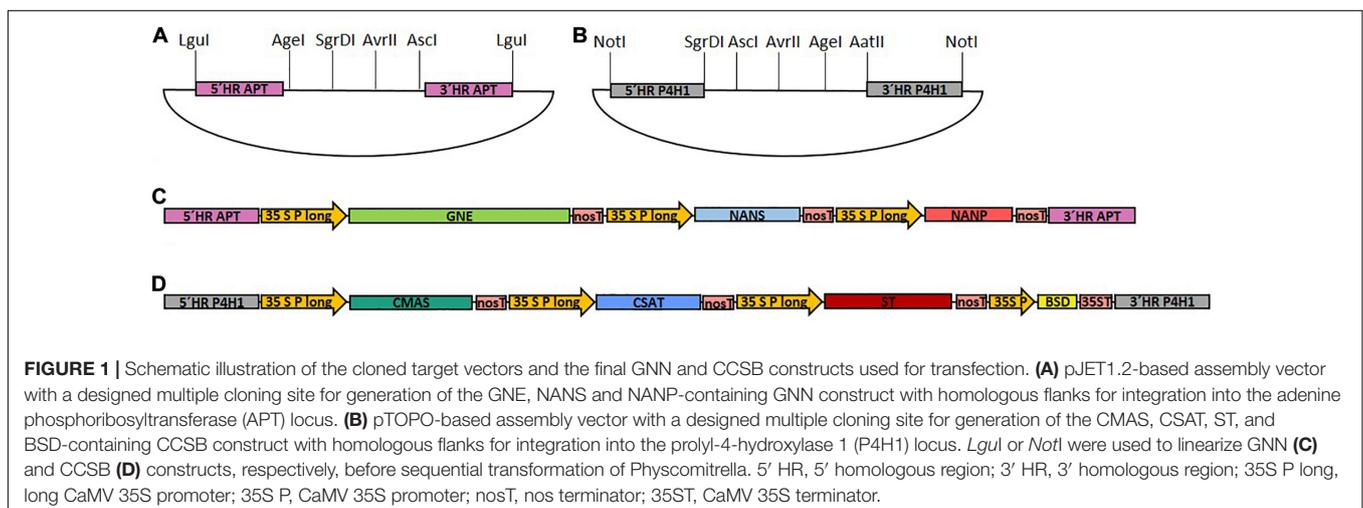
or restriction digest step, agarose gel purification of respective fragments was performed using QIAEX II Gel Extraction Kit (QIAGEN, Hilden, Germany) and ligations were performed using pJET 1.2 Cloning Kit or TOPO™ TA Cloning™ (Thermo Fisher Scientific), all according to the manufacturer's protocols. Assembled vectors were verified by sequencing.

For the heterologous expression of all six genes of the sialylation pathway two multi-gene constructs, containing three PCR-amplified CDS each, were assembled via restriction-site based cloning. GNE (NM_001190414.2) and CSAT (NM_011895.3) were amplified from murine cDNA and ST (NM_001113344.1) from rat cDNA. NANS (NM_018946.4), a truncated CMAS (NM_018686.5) version lacking 120 bases at the 5' end (Castilho et al., 2010) and NANP (NM_152667.3) were amplified from human cDNA. Additionally, a Blasticidin-S deaminase (BSD) expression cassette under the control of the CaMV 35S promoter and CaMV 35S terminator, provides resistance toward Blasticidin S (Kubo et al., 2013).

To assemble the multi-gene constructs, each CDS as well as the BSD cassette were amplified with restriction site-introducing primers (primers 1–14), and subcloned in the pJET 1.2 cloning vector. Expression of each glycosylation-related CDS was driven by the long CaMV 35S promoter (Horstmann et al., 2004) and the nos terminator. Promoter, multiple cloning site (MCS) and terminator were amplified via CDS-specific restriction site-introducing primers (primers 15–22) and subcloned into pJET 1.2, resulting in six different target vectors. To assemble the expression cassettes, the subcloned CDS and the multiple cloning site of the corresponding target vector were digested with the respective restriction enzymes and ligated, resulting in seven expression cassettes, containing GNE, NANS, NANP, CMAS, CSAT, ST, or BSD, respectively. For generation of the two multi-gene constructs, two assembly vectors, either based on pJET 1.2 or pCR™4-TOPO® TA-vector (pTOPO) harboring the corresponding homologous flanks for targeted genome integration and a designed MCS were created, respectively (Figures 1A,B). For the introduction of the first part of the sialylation pathway containing GNE, NANS and NANP

(GNN) vector, homologous flanks were designed to target the integration of the construct into the Physcomitrella adenine phosphoribosyltransferase (APT) gene (Pp3c8_16590V3.1). To create the multiple cloning site of the GNN-assembly vector, the APT-5' homologous flank was amplified with a primer pair introducing an *LguI* restriction site at the 5' end and *AgeI*, *SgrDI* and *AvrII* restriction sites at the 3' end of the PCR product (primers 23 and 24). The APT-3' homologous flank was amplified with a primer pair introducing *AvrII* and *AscI* restriction sites at the 5' end and an *LguI* restriction site at the 3' end of the PCR product (primers 25 and 26). The respective PCR-products were digested with *AvrII*, ligated and subsequently cloned into pJET 1.2, resulting in the GNN assembly vector (Figure 1A). GNE-, NANS- and NANP expression cassettes were subsequently introduced into the GNN-assembly vector using the *AgeI* + *SgrDI*, *SgrDI* + *AvrII* or *AvrII* + *AscI* restriction sites, respectively, resulting in the GNN construct (Supplementary Figure 1B).

For construction of the CMAS, CSAT, ST, and BSD-cassette containing multi-gene construct (CCSB) a similar strategy was used. This construct was created based on the pTOPO vector, targeted to the endogenous prolyl-4-hydroxylase1 (P4H1) gene (Pp3c8_7140V3.1) and additionally harbored a Blasticidin S resistance (BSD) for selection purposes. To create the multiple cloning site of the CCSB-assembly vector, the P4H1-5' homologous flank was amplified with a primer pair introducing a *NotI* restriction site at the 5' end and *SgrDI*, *AgeI* and *AvrII* restriction sites at the 3' end of the PCR product (primers 27 and 28). The P4H1-3' homologous flank was amplified with a primer pair introducing *AvrII*, *AgeI* and *AatII* restriction sites at the 5' end and a *NotI* restriction site at the 3' end of the PCR product (primers 29 and 30). The respective PCR-products were digested with *AvrII*, ligated, polyadenylated with Taq-polymerase using primers 27 and 30 and cloned into the pTOPO vector to create the CCSB-assembly vector (Figure 1B). CMAS, CAST, ST and BSD expression cassettes were subsequently introduced into the CCSB-assembly vector using the *SgrDI* + *AscI*, *AscI* + *AvrII*, *AvrII* + *AgeI*



and *AgeI* + *AatII* restriction sites, respectively, resulting in the CCSB construct (**Supplementary Figure 1C**). The *LguI* or *NotI*-cut GNN and CCSB constructs used for transfections are schematically illustrated in **Figures 1C,D**. A schematic overview of all GNN and CCSB cloning steps is provided in **Supplementary Figure 1**.

GNE Mutation

To circumvent CMP-Neu5Ac-triggered negative feedback regulation of GNE, a mutated version (GNE_{mut}) with amino acid exchanges R263L and R266Q in the allosteric CMP-Neu5Ac binding site (Viswanathan et al., 2005; Son et al., 2011), was created. The corresponding GNE sequence alterations were performed on the previously described GNE expression construct via site-directed mutagenesis using the Phusion Site-Directed Mutagenesis Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. For complete vector amplification a mismatching primer pair, which changes the codon at position 263 from CGA to CTA and replaces the codon CGG at position 266 with CAG, was used (primers 31 and 32). After successful mutagenesis, the native GNE of the GNN construct was replaced by GNE_{mut} via *AgeI* and *SgrDI* restriction sites, resulting in the GM construct.

Cloning of a Chimeric β 1,4-Galactosyltransferase

For cloning of a chimeric β 1,4-galactosyltransferase (FTGT), the localization-determining cytoplasmic, transmembrane, and stem region (CTS) of the human β 1,4-galactosyltransferase (GalT4, NM_001497.4) was replaced by 390 bp encoding the 130 amino acids comprising the CTS of *Physcomitrella* α 1,4-fucosyltransferase (FT4, Pp3c18_90V3.1), responsible for the last step in N-glycosylation. The CTS from the FT4 (FT-CTS) was amplified from moss cDNA with the primers 33 and 34, generating an overlapping region to the GalT4 sequence. In parallel, 927 bp comprising the catalytic domain from the GalT4 CDS were amplified with primers 35 and 36. Subsequently, the PCR products were assembled in a two-template PCR with primers 33 and 36, giving rise to the CDS of the chimeric FT-CTS: β 1,4-galactosyltransferase (FTGT). For the expression of the FTGT, a bicistronic expression cassette driven by the long CaMV 35S promoter was cloned. The sequence of the chimeric FTGT and the *ble* gene (Parsons et al., 2013), conferring resistance to ZeocinTM (Zeocin), were separated by the sequence of a self-cleaving 2A peptide (P2A). This sequence triggers ribosomal skipping during translation, leading to the release of two individual proteins (Donnelly et al., 2001; Kim et al., 2011). P2A was shown to be more efficient than other 2A peptides (Kim et al., 2011). The sequence for P2A was included at the 3' end of the chimeric FTGT CDS by two successive PCR reactions using primers 33 and 37, as well as 33 and 38. The CaMV 35S promoter from the Zeocin resistance cassette in the β 1,3-galactosyltransferase 1-KO construct (Parsons et al., 2012) was excised with *BamHI* and *XhoI* and replaced by a long CaMV 35S promoter amplified with primers 39 and 40 using the same restriction sites. This construct bears homologous flanks targeting the endogenous β 1,3-galactosyltransferase 1 locus (GalT3, Pp3c22_470V3.1) to knock it out (**Supplementary**

Figure 2). The FTGT-P2A sequence was then inserted between promoter and *ble* gene using the *XhoI* site. For transfection the construct was cut using *EcoRI* endonuclease.

Cloning of a Chimeric α 2,6-Sialyltransferase

For the introduction of a chimeric sialyltransferase (FTST), consisting of the endogenous FT-CTS and the catalytic domain of the rat α 2,6-sialyltransferase, a construct with homologous flanks for targeted integration into the P4H2 locus (Pp3c20_10350V3.1) (**Supplementary Figure 3**) was cloned via Gibson assembly. Amplification of the 5' and 3' homologous flanks was performed from moss genomic DNA using overhang introducing primers 41 and 42 or 43 and 44, respectively. The hygromycin selection cassette coding for a hygromycin B phosphotransferase (hpt) (Decker et al., 2015) was amplified using the primers 45 and 46. Gibson assembly was performed as described in Gibson et al. (2009) into a pJET 1.2 backbone. For introduction of the chimeric FTST variant this vector was linearized with *BseRI* at the 3' end of the 5' homologous flank and subsequently used as backbone for the introduction of the FTST anew via Gibson assembly. For this the FT-CTS together with the long CaMV 35S promoter was amplified with the primers 47 and 48 from the previously described FTGT construct and the catalytic domain of the rat ST including the nos terminator was amplified from the previously described CCSB construct using primers 49 and 50. For transfection the construct was cut using *NotI* and *XbaI* endonucleases.

Plant Material and Cell Culture

Physcomitrella was cultivated as described previously (Frank et al., 2005) in Knop medium supplemented with microelements (KnopME, Horst et al., 2016). The sialylating lines were obtained by stable transformation of the Δ xt/ft moss line (IMSC no.: 40828), in which the α 1,3-fucosyltransferase and the β 1,2-xylosyltransferase genes have been disrupted via homologous recombination (Koprivova et al., 2004). This line additionally produces a recombinant human reporter glycoprotein.

Transfection of *Physcomitrella* Protoplasts and Selection of Transgenic Lines

Protoplast isolation, transformation and regeneration were performed as described before (Decker et al., 2015). For each transfection event 50 μ g of linearized plasmid were used. Selection for the marker-free GNN or GM lines with targeted ATP locus disruption was performed on 0.3 mM 2-fluoroadenine (2-FA) containing solid KnopME plates. Protoplasts were regenerated in liquid regeneration medium for 5 days, then transferred to KnopME solid medium covered with cellophane for 3 days. Subsequently, the cellophane sheets with the regenerating protoplasts were transferred onto 2-FA containing KnopME plates for a further 3 weeks. Selection with Blasticidin S (Sigma-Aldrich) or Zeocin (Invitrogen) were started in liquid regeneration medium on day 8 after transfection via the addition of 75 mg/L Blasticidin S or 50 mg/L Zeocin, respectively. After 4 days of selection in liquid medium the protoplasts were transferred to solid KnopME plates containing either 75 mg/L Blasticidin S or 100 mg/L Zeocin and 1% MES, respectively. The

selection was done in two successive cycles of 3 weeks on selective plates interrupted by a 2-week release on non-selective KnopME plates. Hygromycin selection was performed as described before (Wiedemann et al., 2018).

Molecular Validation of Transgenic Moss Lines

Plants surviving the selection were screened for targeted construct integration via direct PCR (Schween et al., 2003). Successful extraction of DNA, which was performed as described before (Parsons et al., 2012), was assayed by amplifying a part of the elongation factor 1 gene (EF1 α , Pp3c2_10310V3.1) with primer pair 51 and 52. Targeted integration of the GNN or GM constructs in the APT locus was confirmed using the primer pairs 53 and 54 for 5'- and 55 and 56 for 3'-integration. Targeted integration of the CCSB construct within the P4H1 locus was assayed with the primers 57 and 58 as well as 59 and 60 for the 5'- and 3'-integration, respectively. Homologous integration of the chimeric β 1,4-galactosyltransferase FTGT into the GalT3 locus was confirmed using the primers 61 and 62 (5'-integration), whereas correct 3'-integration was confirmed by the primers 63 and 64. Targeted integration of the FTST construct into the P4H2 locus was assayed with the primers 65 and 66 for 5'- and 67 and 68 for 3'-integration.

Expression Analysis of Transgenic Moss Lines

Real-Time qPCR (qRT-PCR) Analysis of Gene Expression

To determine the expression levels of the transgenes, total RNA was isolated using Trizol (Thermo Fischer Scientific) according to the manufacturer's instructions. RNA concentrations were determined via UV-Vis spectrometric measurement (NanoDrop ND 1000, PEQLAB Biotechnologie GmbH, Erlangen) and quality was checked via agarose gel electrophoresis. Isolated RNA was subsequently digested with DNaseI (Thermo Fisher Scientific) and cDNA synthesis was performed with random hexamers using TaqMan[®] Reverse Transcription Reagents (Thermo Fisher Scientific), both according to the manufacturer's protocols. Completeness of DNaseI digestion was confirmed by a non-transcribed control, without the addition of MultiScribe[®] RT enzyme. Primer pairs were designed using the Universal Probe Library by Roche¹ and selected according to the lowest amount of off-target hits identified in a Phytozome (Goodstein et al., 2012) search against the Physcomitrella transcriptome (V3.3; Lang et al., 2018). Oligonucleotide pair efficiencies of 2 were confirmed prior to analysis with a qPCR-run of a serial 1:2 dilution row of cDNA completed by a control without cDNA. Presence of off-targets was excluded via melting curve analysis. Measurements were prepared in white 96-multiwell plates and carried out in triplicate, using the SensiMix[™] Kit and SYBRGreen (Bioline, Luckenwalde, Germany). For each triplicate cDNA amounts corresponding to 50 ng RNA were used and measurements

were conducted in a LightCycler[®] 480 (Roche) according to the manufacturer's instructions. Absence of DNA contamination in reagents was confirmed for each primer pair with a control reaction without template addition. Amplification was performed in 40 cycles with a melting temperature of 60°C. Computational analysis of the qRT-PCR was done with LightCycler[®] 480 software (Roche). The analysis of the expression levels of the transgenes was performed relatively against the expression values of the internal housekeeping genes coding for EF1 α and for the ribosomal protein L21 (Pp3c13_2360V3.1) (Beike et al., 2015). The relative expression of the gene of interest (GOI) compared to the controls (ctrl) EF1 α and L21 was calculated as $2^{(-\Delta\Delta C_T)}$, assuming an primer efficiency of 2, and where $\Delta C_T = C_{TGOI} - C_{Tctrl}$ and C_T is defined as the cycle number at which each sample reaches an arbitrary threshold (Livak and Schmittgen, 2001). Primer pairs used: GNE: 73 and 74; NANS: 75 and 76; NANP: 77 and 78; CMAS: 79 and 80; CSAT: 81 and 82; ST and FTST: 83 and 84; FTGT: 85 and 86; EF1 α : 69 and 70 and L21: 71 and 72.

RNAseq Library Preparation and Data Analysis

Total RNA was extracted from 100 mg fresh weight (FW) protonema tissue in biological triplicates for each line, using Direct-zol[™] RNA MicroPrep Kit (Zymo Research, Freiburg) according to the manufacturer's protocol. RNAseq library preparations were performed by the Genomics Unit at Instituto Gulbenkian de Ciencia (Portugal) with conditions optimized according to Picelli et al. (2014); Baym et al. (2015), and Macaulay et al. (2016) using Smart-seq2. Sequencing was performed on an Illumina NextSeq 500 instrument producing 75 bp long single-end reads.

We assessed sequence quality with FastQC (Galaxy Version 0.72 + galaxy1; Andrews, 2010) and MultiQC (Galaxy Version 1.6; Ewels et al., 2016) on the public European Galaxy instance at <https://usegalaxy.org> (Afgan et al., 2018). For transcript quantification, reads of each library were pseudoaligned against the Physcomitrella transcriptome (V3.3; Lang et al., 2018), obtained from Phytozome v12.1.5 (Goodstein et al., 2012) and complemented with the sequences of the introduced transgenes using Kallisto quant (Galaxy Version 0.43.1.4; Bray et al., 2016) for single-end reads in "unstranded" mode. According to the information obtained from the sequencing facility, the average fragment length of the libraries was specified as 340 with an estimated standard deviation of 34. For the quantification algorithm a number of 100 bootstrap samples was chosen. The data used for these analyses can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA665456.

Detection of Sialic Acid (Neu5Ac) via Periodate-Resorcinol-Assay

The detection of total sialic acid concentrations in moss extracts was performed according to Jourdian et al. (1971) with the following modifications: Protonema, the young filamentous tissue of Physcomitrella, was harvested via vacuum filtration and frozen. Approximately 150 mg FW of frozen plant material was disrupted with a glass and a metal bead (\varnothing 3 mm; QIAGEN GmbH, Hilden, Germany) in a TissueLyser (MM400, Retsch

¹https://lifescience.roche.com/en_de/brands/universal-probe-library.html#assay-design-center

GmbH, Haan, Germany) at 30 Hz for 1.5 min. For Neu5Ac extraction, the fourfold amount (v/w) of 50 mM Tris/HCl (pH 7.0) was added and the samples were vortexed for 10 min and afterward sonicated (Bandelin Sonorex RK52, Bandelin electronic GmbH & Co. KG, Berlin, Germany) for 15 min at 4°C. Crude extracts were cleared via two subsequent centrifugation steps at 14,000 rpm and 4°C for 10 min and further 30 min centrifugation of the supernatant in a fresh reaction tube. Serial dilutions of the samples (1:2 to 1:32) and a standard row containing 1 to 40 nmol Neu5Ac (Sigma-Aldrich) were prepared in extraction buffer, final volume 120 µl. For sialic acid oxidation 30 µl of a 0.032 M periodic acid solution were added, followed by 45 min incubation under gently shaking conditions at 4°C. Samples were additionally cleared via a 10 min centrifugation step at 14,000 rpm and 4°C. Subsequently, 100 µl of freshly prepared resorcinol solution (0.06% w/v resorcinol (Sigma-Aldrich), 16.8% HCl, 0.25 µM CuSO₄) and 50 µl sample or standard were mixed in a 96-well plate (Greiner Bio-One, Frickenhausen, Germany), sealed and incubated for 45 min at 80°C. Color complexes were stabilized by the addition of 100 µl tert-butanol (Sigma-Aldrich) and the absorbance measured at 595 nm (Sunrise absorbance reader, Tecan, Männedorf, Switzerland, software Magellan™ V 7.1). Calculation of Neu5Ac concentration in the samples was performed by linear regression.

Detection of Sialic Acid (Neu5Ac) via RP-HPLC-FLD

Detection of sialic acid in moss extracts was performed as described previously for *Arabidopsis thaliana* (Castilho et al., 2008) via reverse-phase high-performance liquid chromatography coupled with fluorescence detection (RP-HPLC-FLD). Moss protonema tissue was dispersed with an ULTRA-TURRAX® (IKA, Staufen, Germany), 1 ml of the crude extract was taken for dry weight (DW) determination and further 500 µl were mixed with 50 µl acetic acid and incubated for 10 min under shaking conditions. Extracts were cleared via 1 min of centrifugation at 16,100 × g and supernatants were further purified via a C18 SPE column (Strata C18-E, 50 mg; Phenomenex), pre-equilibrated with 1% acetic acid. After loading to the column, the samples were washed with 200 µl 1% acetic acid. The column flow-through was vacuum-dried and afterward resuspended in 30 µl ultra-pure water. Ten microliter of the C18 SPE purified extracts were DMB-labeled with 60 µl of DMB labeling reagent (Sigma-Aldrich), at 50°C for 2.5 h in the dark under shaking conditions at 750 rpm (Hara et al., 1987). Prior to analysis, quenching of the reaction was performed by the addition of 730 µl of ultra-pure water. For detection of sialic acid, 5 to 20 µl of the labeled extracts were injected on a HPLC (Nexera X2 HPLC system) with a RF-20Axs Fluorescence Detector, equipped with a semimicro flow cell (Shimadzu, Korneuburg, Austria). Separation was performed on an Aquasil C18 column (250 cm × 3 mm, 5 µm particle size; Thermo Fisher Scientific) at a flow rate of 1 ml/min, applying a linear gradient from 30% to 42% B (70% 100 mM ammonium acetate pH 5.5, 30% acetonitrile; Solvent A was water) over 12 min and the column thermostat

was set to 35°C. Fluorescence was measured with wavelengths excitation/emission 373 nm and 448 nm. Identification of DMB-Neu5Ac in the moss extract was performed in comparison to a DMB-Neu5Ac standard sample and confirmed by its co-eluting fluorescent profile; its concentration was determined via peak area integration.

Detection of Activated Sialic Acid (CMP-Neu5Ac) via Mass Spectrometry

Five ml of protonema suspension culture were supplemented to an ammonia concentration of 1% (v/v) and dispersed with an ULTRA-TURRAX®. Afterward the extracts were cleared via 10 min of centrifugation at 16,100 × g and supernatants were applied on a 10 mg HyperSep™ Hypercarb™ SPE cartridge (Thermo Fisher Scientific). Washing of the column was conducted with 1 ml 1% ammonia, and CMP-Neu5Ac was eluted with 600 µl 50% ACN in 1% ammonia. To minimize degradation of CMP-Neu5Ac, this extraction process was performed in the minimal possible time (max. 30 min). In parallel, degradation of a thawed and frozen CMP-Neu5Ac standard (Sigma-Aldrich) was assayed, to exclude that CMP-Neu5Ac degradation compromise the results. Afterward, the elution fraction was vacuum-dried and resuspended in 10 µl 80 mM formic acid, buffered to pH 9 with ammonia. Five microliter were injected on a Dionex Ultimate 3000 LC-system, using a Hypercarb™ Porous Graphitic Carbon LC Column (150 × 0.32 mm; Thermo Fisher Scientific). Solvent A consisted of 80 mM formic acid, buffered to pH 9, solvent B of 80% ACN in solvent A. At a flow of 6 µl/min and column oven set to 31°C, initial conditions of 1.3% B were held for 5 min, went to 19% B over 27 min and finally 63% in 1 min, which was held for 7 min. The LC was directly coupled to a Bruker amaZone speed ETD ion trap instrument (Bruker, Bremen, Germany) with standard ESI source settings (capillary voltage 4.5 kV, nebulizer gas pressure 0.5 bar, drying gas 5 L/min, 200°C). Spectra were recorded in negative ion data depended acquisition mode with MS¹ set on *m/z* 613.1 which corresponds to the [M-H]⁻ ion of CMP-Neu5Ac, and simulated selected ion monitoring of *m/z* 322.0 ([M-H]⁻ of CMP) was performed with MS². Confirmation of CMP-Neu5Ac identity was performed via the measurement of a CMP-Neu5Ac standard demonstrating an identical fragment pattern on MS² level. Quantification of CMP-Neu5Ac amounts were performed via peak area integration and compared to peak areas gained from defined amounts of a CMP-Neu5Ac standard.

Glycoprotein Analysis

Protein Extraction for Mass Spectrometric Analyses

Moss protonema material (1–2 g of fresh weight), cultivated for 6 days in KnopME at pH 4.5 with 2.5 mM ammonium tartrate (Sigma-Aldrich) was harvested over a 100 µm sieve and resuspended in threefold amount extraction buffer (408 mM NaCl, 60 mM Na₂HPO₄ × 2H₂O, 10.56 mM KH₂PO₄, 60 mM EDTA, 1% protease inhibitor (Sigma-Aldrich), pH 7.4). The material was homogenized with an ULTRA-TURRAX® for 10 min at 10,000 rpm on ice. The crude cell lysate was cleared via two successive centrifugation steps for 5 min at 5,000 rpm and

4°C. Extracts were frozen in 150 µL aliquots, in liquid nitrogen and stored at -80°C.

Glycopeptide Analysis

For mass spectrometry (MS) protein extracts were supplemented with 2% SDS and 25 mM DTT (final concentration) and incubated for 10 min at 90°C. After cooling to room temperature, proteins were S-alkylated with 60 mM iodoacetamide for 20 min in darkness. Prior to SDS-PAGE the samples were mixed with 4× sample loading buffer (Bio-Rad, Munich, Germany). Separation of proteins was carried out via SDS-PAGE in 7.5% gels (Ready Gel Tris-HCl; BioRad) in TGS buffer (BioRad) at 120 V. After electrophoresis, the gel was washed three times for 10 min in water followed by a 1 h staining period with PageBlue® Protein Staining Solution (Thermo Fisher Scientific). For MS analysis gel bands corresponding to high molecular weight range proteins were excised. Gel band preparations, MS measurements on a QExactive Plus instrument and data analysis were performed as described previously (Top et al., 2019). Raw data were processed using Mascot Distiller V2.5.1.0 (Matrix Science, United States) and the peak lists obtained were searched with Mascot V2.6.0 against an in-house database containing all *Physcomitrella* V3.3 protein models (Lang et al., 2018) and the recombinant reporter protein. Glycopeptides were identified from the Mascot mgf files using a custom Perl script. Glycopeptide precursor masses were searched within these files with a precursor mass tolerance of ±5 ppm and further validated by the presence of typical GlcNAc oxonium ions (m/z -values: $[\text{GlcNAc}]^+ = 204.087$, $[\text{GlcNAc} - \text{H}_2\text{O}]^+ = 186.076$, $[\text{GlcNAc} - 2\text{H}_2\text{O}]^+ = 168.066$, $[\text{GlcNAc} - \text{C}_2\text{H}_4\text{O}_2]^+ = 144.065$, $[\text{GlcNAc} - \text{CH}_6\text{O}_3]^+ = 138.055$, $[\text{GlcNAc} - \text{C}_2\text{H}_6\text{O}_3]^+ = 126.055$), glycan fragment ions ($[\text{GlcNAcHex}]^+ = 366.139$, $[\text{GlcNAcHex}_2]^+ = 528.191$) or in the case of sialic acid containing glycopeptides the presence of Neu5Ac-specific oxonium ions (m/z -values: $[\text{Neu5Ac}]^+ = 292.103$ and $[\text{Neu5Ac} - \text{H}_2\text{O}]^+ = 274.092$) (Halim et al., 2014). A fragment mass tolerance of 0.02 Da was used. A list of the glycopeptides searched and their calculated precursor masses is provided in **Supplementary Table 2**. Quantification of identified glycopeptides was done using a default MaxQuant search (V1.6.0.16) on the raw data. For each glycopeptide identified from Mascot mgf files, the intensity value (peak area) was extracted from the MaxQuant allPeptides.txt file by matching raw file name, precursor mass and MS² scan number using another custom Perl script.

RESULTS

Generation of Stable Sialic Acid-Producing Lines

To establish protein sialylation in *Physcomitrella*, six sequences encoding the enzymes required to synthesize, activate, transport and finally transfer sialic acid onto β1,4-galactosylated N-glycans needed to be introduced into the moss genome. These CDSs were distributed to two multi-gene cassettes, with homologous flanks for targeted integration into the *Physcomitrella* genome

(**Figures 1C,D**), of a Δxt/ft (β1,2-xylosyltransferase and the α1,3-fucosyltransferase) knock-out line expressing a recombinant human reporter glycoprotein to verify the N-glycan sialylation.

As plants have no or undetectable amounts of ManNAc (Paccalet et al., 2007), synthesis of sialic acid in plants needs to be started with the epimerization of UDP-GlcNAc by GNE. To enable the production of sialic acid, the enzymes responsible for its biosynthesis, GNE, NANS and NANP were introduced via two versions of a multi-gene expression construct only differing in the GNE CDS used. One expression construct (GNN) contained the native GNE and in the other the GNE CDS was altered with two point mutations, R263L and R266Q, in the CMP-Neu5Ac binding site (GM). Both constructs bear homologous regions targeting the adenine phosphoribosyltransferase locus, whose disruption leads to interruption of the adenine-salvage-pathway resulting in lines resistant toward the toxic adenine analog 2-fluoro adenine (2-FA) (Trouiller et al., 2007). The selection on 2-FA resulted in four GNN-lines with the native GNE (GNN1-4) and 64 GM-lines with the mutated GNE, which were directly screened by PCR for targeted integration of the respective construct within the genome. Homologous 5'- and 3'-construct integration was confirmed for two lines harboring the GNN construct (GNN2 and GNN4) and 3' homologous integration was confirmed for two further GNN-lines (GNN1 and GNN3) as well as 26 GM-construct containing lines (GM9, 10–12, 16, 21, 22, 25–28, 36, 41, 44–47, 49, 51, 52, 57, and 61–65), respectively (**Supplementary Figures 4, 5**). These lines were chosen for further experiments.

GNN and GM Lines Produce Sialic Acid

In mammals, the coordinated activity of GNE, NANS, and NANP leads to the synthesis of free sialic acid. To prove the activity of these enzymes in *Physcomitrella*, sialic acid was quantified in plant extracts via a colorimetric periodate-resorcinol assay, which measures total sialic acid. Three out of four analyzed GNN-lines and 20 out of 22 analyzed GM-lines could be identified as sialic acid producers, while, as expected, the parental line does not show any signal indicating sialic acid production. Sialic acid levels between 9.8 ± 0.5 and 16.5 ± 0.7 µmol/g FW were detected in the GNN lines, whereas similar amounts ranging from 4.1 ± 0.1 up to 15.9 ± 0.5 µmol/g FW were obtained in the GM lines (**Figure 2A**). This indicates, on the one hand, that the epimerase and kinase activities of the mutated GNE are not affected, as was already shown for this mutated GNE version in CHO cells (Son et al., 2011). Moreover, the presence of free sialic acid as product of the reaction of GNE, NANS and NANP, confirms the sequential activity of the three enzymes. These include the synthesis of ManNAc and subsequently ManNAc-6-P out of UDP-GlcNAc catalyzed by GNE, the NANS mediated condensation of the ManNAc-6-P with PEP resulting in Neu5Ac-9-P and finally the formation of Neu5Ac catalyzed by NANP. For the overall best producing line GNN2 presence of sialic acid was confirmed via fluorescence detection of DMB-labeled moss extracts separated over reverse-phase HPLC (RP-HPLC-FLD) compared to a DMB-Neu5Ac standard and the parental line as a negative control (**Figure 2B** and **Supplementary Figure 6**). According to the peaks' area ratio, 85 µmol Neu5Ac/g DW could be detected in GNN2 extracts. Considering that under

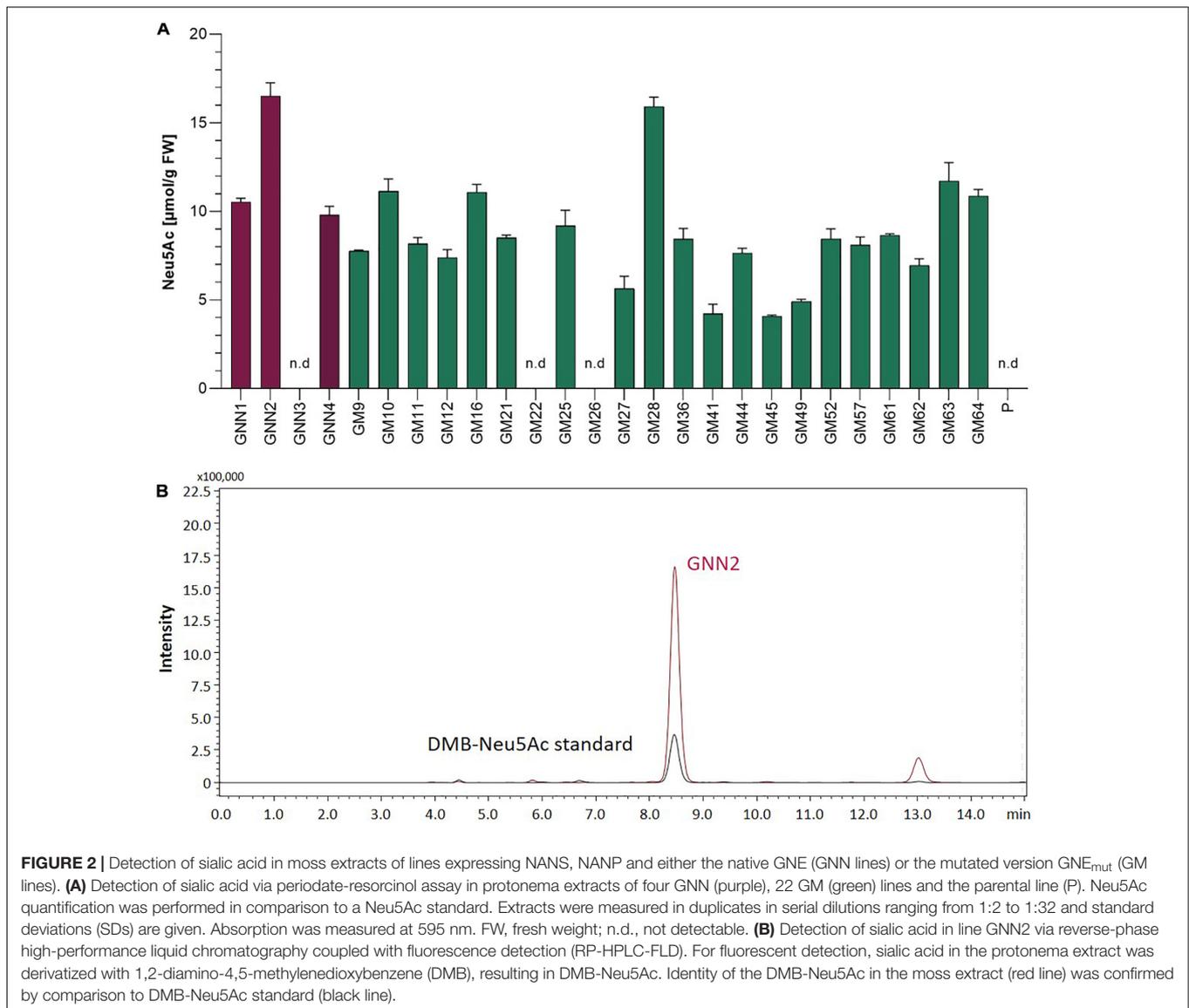


FIGURE 2 | Detection of sialic acid in moss extracts of lines expressing NANS, NANP and either the native GNE (GNN lines) or the mutated version GNE_{mut} (GM lines). **(A)** Detection of sialic acid via periodate-resorcinol assay in protonema extracts of four GNN (purple), 22 GM (green) lines and the parental line (P). Neu5Ac quantification was performed in comparison to a Neu5Ac standard. Extracts were measured in duplicates in serial dilutions ranging from 1:2 to 1:32 and standard deviations (SDs) are given. Absorption was measured at 595 nm. FW, fresh weight; n.d., not detectable. **(B)** Detection of sialic acid in line GNN2 via reverse-phase high-performance liquid chromatography coupled with fluorescence detection (RP-HPLC-FLD). For fluorescent detection, sialic acid in the protonema extract was derivatized with 1,2-diamino-4,5-methylenedioxybenzene (DMB), resulting in DMB-Neu5Ac. Identity of the DMB-Neu5Ac in the moss extract (red line) was confirmed by comparison to DMB-Neu5Ac standard (black line).

our experimental conditions the relationship FW:DW is 5:1, this result is in agreement with the result of the colorimetric assay.

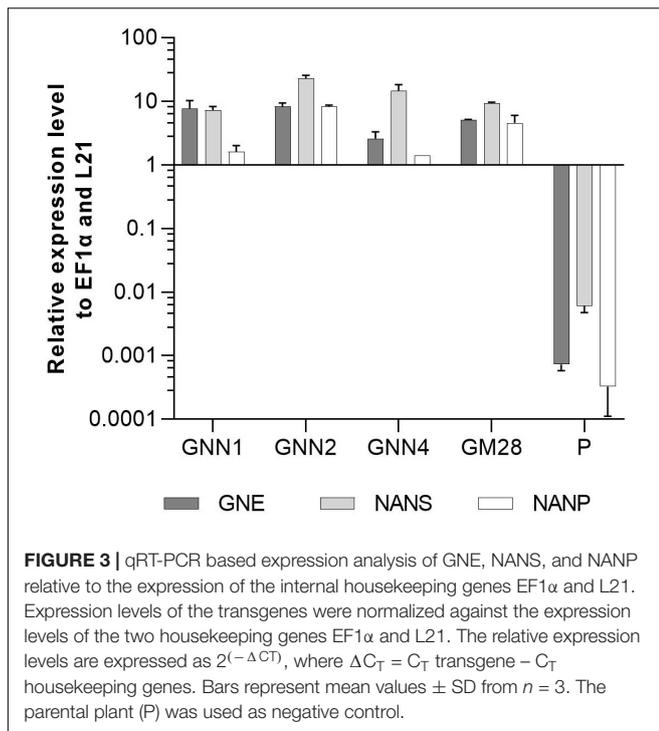
Expression Analysis of GNE, NANS, and NANP in GNN and GM Lines

Four Neu5Ac-producing lines (GNN1, GNN2, GNN4 and GM28) were characterized via Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) regarding their transgene expression levels. The results are summarized in **Figure 3** via the respective $2^{(-\Delta\text{CT})}$ -values which represent the ratios between the expression level of the respective transgene compared to the expression levels of the housekeeping genes EF1 α and L21. All analyzed lines strongly express the introduced transgenes. The parental line, expressing the recombinant reporter protein but none of the sialylation-pathway genes, was used as negative control. Therefore the values obtained for the sialylation-pathway genes in the control were considered as background

signal (**Figure 3**). The expression level of the transgenes varied in the different lines, notably the expression level of NANP seems to be related to the amount of free sialic acid found in these lines (**Figures 2A, 3**), indicating that the expression of this enzyme might play an important role in the efficiency of sialic acid synthesis.

Generation of the Complete Sialylation Pathway

After confirming the synthesis of sialic acid, the second part of the sialylation pathway catalyzing the activation, transport and glycosidic transfer of sialic acid was introduced. For this, the best Neu5Ac producing lines GNN2 and GM28 were transfected with the second multi-gene construct, called CCSB, containing the CDSs of CMAS, CSAT, ST as well as a Blasticidin S resistance cassette targeted to the moss P4H1 gene, to prevent plant typical prolyl hydroxylation of recombinant



proteins (Parsons et al., 2013). After selection, surviving lines were directly screened by PCR for targeted integration of the CCSB construct into the P4H1 locus. This resulted in two lines in the GNN2-background (GNC lines 4 and 7, **Supplementary Figure 7**) and four lines with confirmed 3'-integration in the GM28 background (GMC lines 5, 22, 23, and 46, **Supplementary Figure 8**).

Expression Analysis of CMAS, CSAT, and ST

The expression levels of CMAS, CSAT, and ST were analyzed via qRT-PCR in all six lines with targeted integration of the CCSB construct alongside their respective parental lines GNN2 for GNC and GM28 for GMC as negative controls. The expression levels of previously introduced GNE, NANS and NANP were analyzed anew. In four of these lines, GNC7, GMC5, GMC23, and GMC46, all six transgenes are strongly expressed, ranging from comparable to higher expression levels than the strongly expressed housekeeping genes EF1 α and L21 (**Figure 4**). The other two lines showed a low expression level for at least one gene, mainly the CMAS. Unexpectedly, in GNC4 a marked decrease in the expression levels of GNE, NANS and NANP was observed.

Mass Spectrometric Determination of CMP-Neu5Ac in Lines Expressing the Complete Sialylation Pathway

Activation of the free sialic acid by addition of CMP is a prerequisite for later transfer to the N-glycan structure. Therefore, the four lines robustly expressing all six transgenes (GNC7, GMC5, GMC23, and GMC46) were tested regarding their ability to activate sialic acid by measuring the CMP-Neu5Ac content via mass spectrometry in negative ion mode. The expected [CMP-Neu5Ac - H]⁻ at m/z 613.1 on MS¹ level and the CMP-fragment ion [CMP - H]⁻ at m/z 322.0 on MS² level

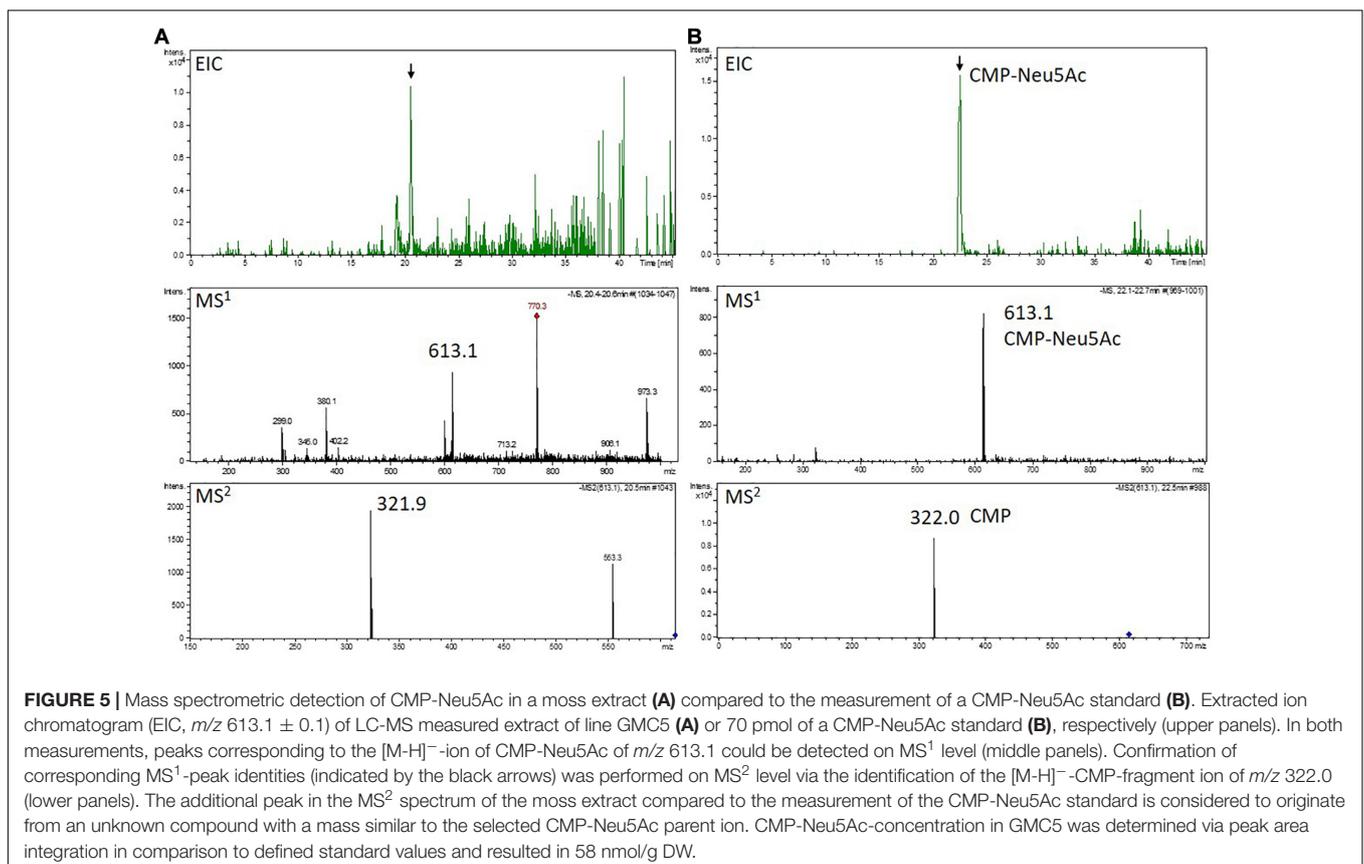
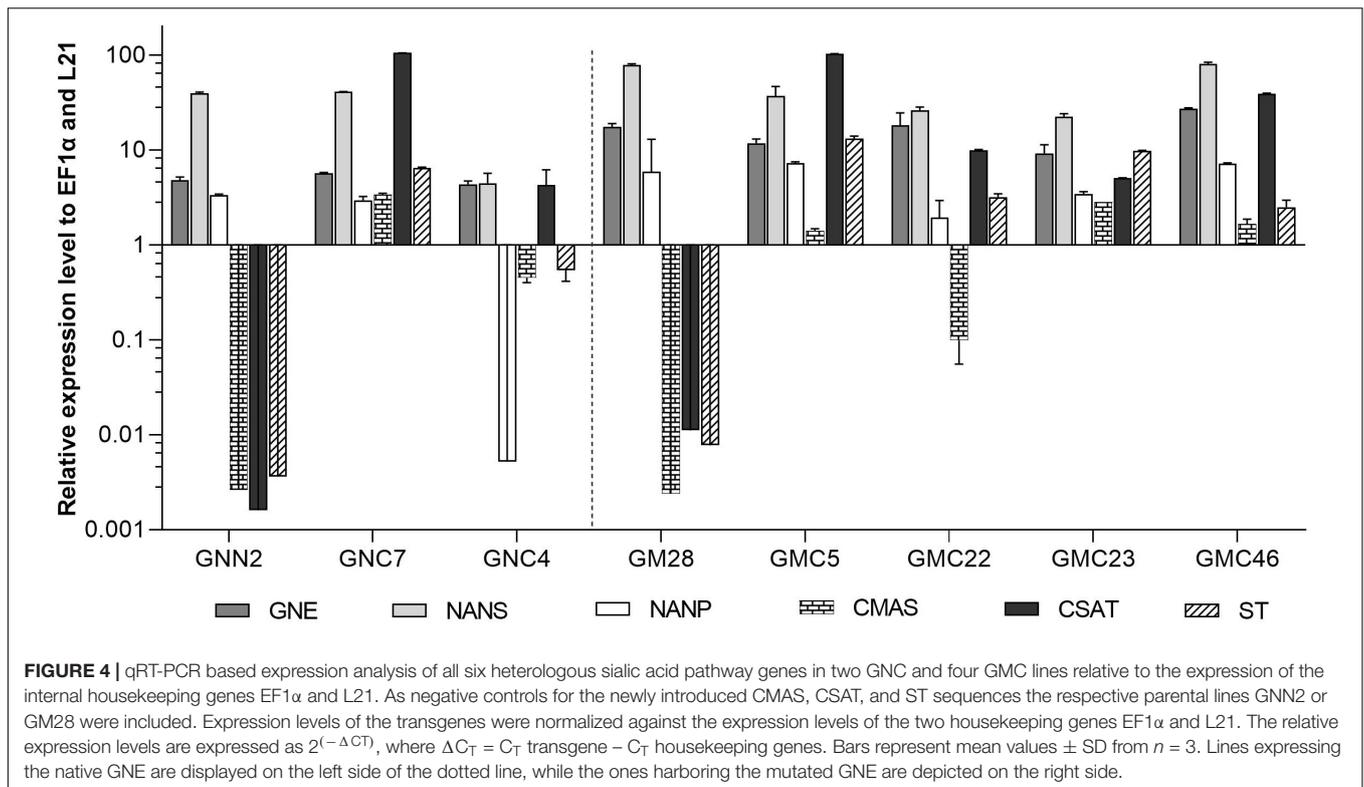
could be detected in all analyzed lines, confirming the presence of CMP-Neu5Ac exemplarily shown for GMC5 in **Figure 5**. Verification of CMP-Neu5Ac presence in the other lines as well as a measurement of a non-CMP-Neu5Ac producing line are depicted in **Supplementary Figure 9**. Peak area integration of the respective extracted ion chromatograms (EICs) corresponding to the m/z -value of 613.1 in comparison to the standard yielded in CMP-Neu5Ac amounts of only 2 nmol/g DW in line GNC7, in contrast to 14 nmol/g DW in the lines GMC23 and GMC46 and 58 nmol/g DW in the line GMC5. This confirmed the activity of the CMAS version used in moss leading to the production of activated sialic acid. Additionally, CMP-Neu5Ac-levels in GNE mutated lines were up to 25-fold higher than those in the native GNE-harboring line GNC7.

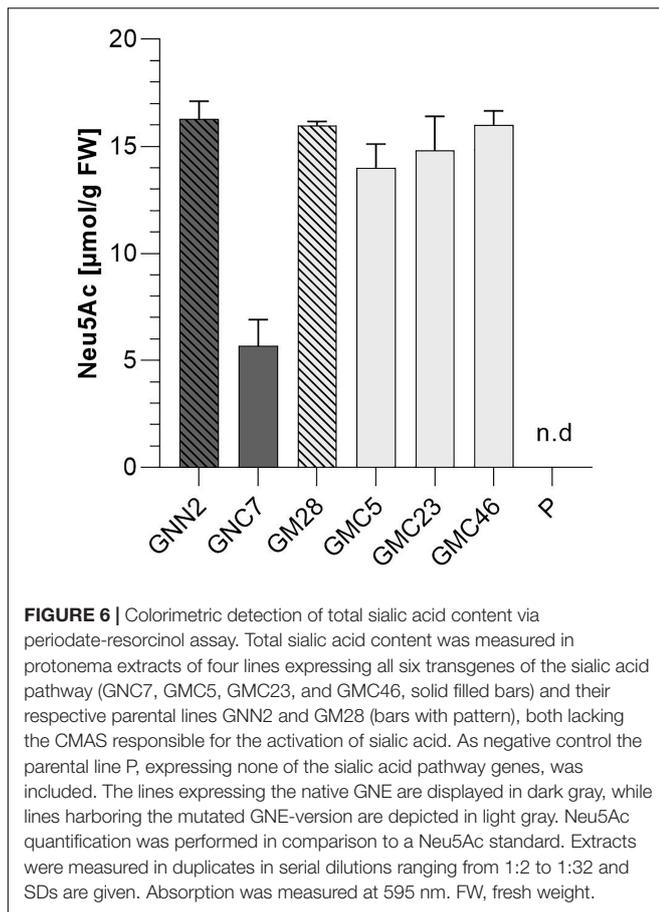
CMP-Neu5Ac Triggers GNE Feedback Inhibition, Which Can Be Overcome by Mutating GNE

Due to the difference in CMP-Neu5Ac amounts between plants harboring the two different versions of GNE, sialic acid content of all four selected lines expressing the six sialylation pathway-related transgenes was quantified again via periodate resorcinol assay in comparison to their respective parental lines GNN2 and GM28 as well as the parental line P, expressing none of the sialic acid pathway genes. This analysis revealed that with the introduction of the second half of the sialylation pathway the Neu5Ac content declined over 60% in the native GNE-expressing line GNC7 compared to the GNN2 parental line (**Figure 6**), despite no changes of the GNE, NANS and NANP expression levels being detected (**Figure 4**). In contrast, in the three analyzed GMC lines 5, 23, and 46, expressing the mutated GNE-version, the sialic acid content remained stable compared to the GM28 parental line, while in the parental line P no sialic acid could be detected (**Figure 6**). The higher content of CMP-Neu5Ac in lines carrying the GNE_{mut} in comparison to the native GNE containing line GNC7 indicates that the negative feedback of the activated sialic acid on the activity of GNE could be overcome with the two point mutations, which was previously described for CHO cells (Son et al., 2011). According to these results the GNE-mutated lines were chosen for further experiments.

Generation of N-Glycan Galactosylation by Introduction of the β 1,4-Galactosyltransferase

After the introduction of the whole sialylation pathway, we aimed to provide terminal β 1,4-linked galactoses for the anchoring of sialic acid to N-glycans. For this, a construct to simultaneously introduce the human β 1,4-galactosyltransferase (GalT4) and to knock out the undesired activity of the endogenous β 1,3-galactosyltransferase 1 (GalT3), responsible for the formation of Le^a epitopes (Parsons et al., 2012), was generated. Moreover, as the glycosyltransferases in the Golgi apparatus should act in a sequential fashion, for appropriate localization of the GalT4 in the late plant Golgi apparatus (Bakker et al., 2006), a chimeric version of the enzyme was designed, bearing the CTS

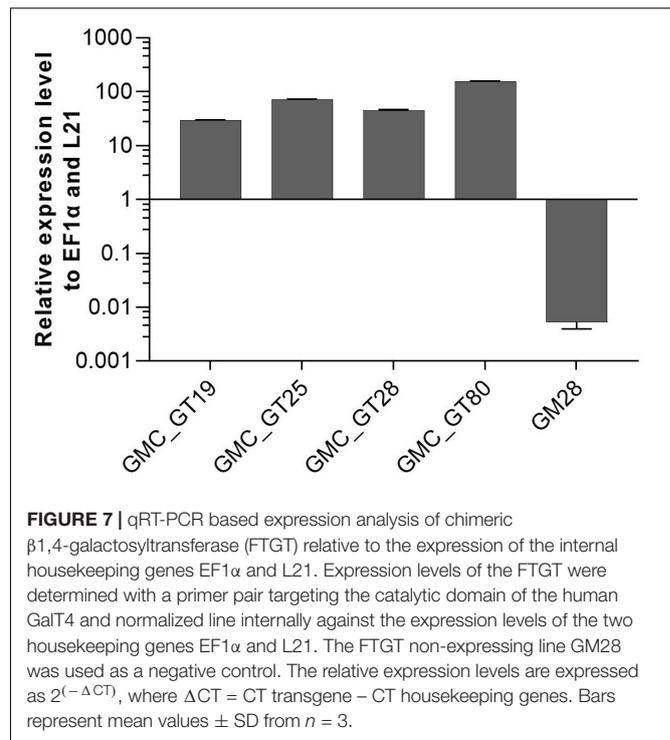




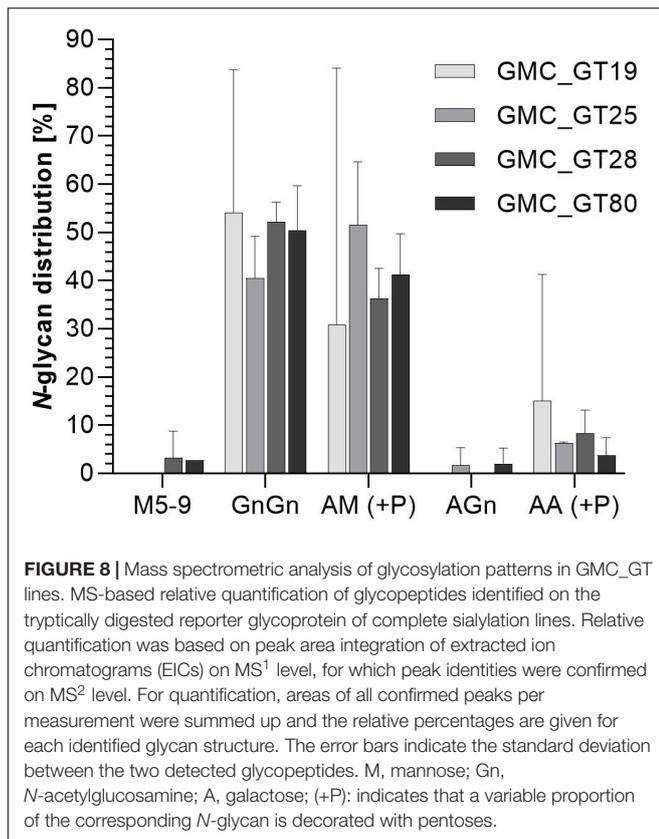
of the last acting enzyme in plant N-glycosylation, the α 1,4-fucosyltransferase. This CTS region was N-terminally fused to the catalytic domain of the GalT4, resulting in the chimeric FT-CTS-GalT4 (FTGT). FTGT expression was driven by the long CaMV 35S promoter, which is according to Horstmann et al. (2004) four times stronger than the CaMV 35S promoter in *Physcomitrella*. Line GMC23, with a high expression level of all six genes introduced previously, with the GNE_{mut} version, was transformed with the FTGT-encoding construct. After Zeocin selection, resulting lines were screened for homologous integration in the GalT3 locus, which resulted in four lines with targeted FTGT-construct integration (GMC_GT19, GMC_GT25, GMC_GT28 and GMC_GT80, **Supplementary Figure 10**). These four lines as well as the line GM28 as a negative control, were analyzed for the respective expression level via qRT-PCR. All FTGT construct-containing lines expressed the chimeric β 1,4-galactosyltransferase strongly, much higher than the housekeeping genes EF1 α and L21, whereas in GM28 no FTGT expression was detected (**Figure 7**).

Mass Spectrometric Analysis of the N-Glycosylation Pattern of Complete Sialylation Lines

The efficiency of galactosylation and subsequent sialylation of N-glycans in the FTGT-expressing lines GMC_GT19, 25, 28, and



80, was assessed on glycopeptides of the stably co-expressed reporter protein via mass spectrometry. Evaluation of the N-glycosylation pattern was performed via integration of MS²-confirmed glycopeptide elution profiles on MS¹ level. Our present glycopeptide analysis does not allow us to distinguish conformational isomers of N-glycans but enables us to make conclusions about their composition. For simplification in the description of the results, we present all possible isomers of an N-glycan combined under one structure, e.g., AM, MA or a mixture of both, are all displayed together as AM. The different glycan structures are depicted in **Supplementary Table 3**. The MS-analyses revealed that the investigated lines displayed a galactosylation efficiency between 45% and 60% within the analyzed glycopeptides (**Figure 8**), confirming the functional activity of the introduced chimeric FTGT variant. Among galactosylated glycopeptides, 80% corresponded to AM-structures. On average, 8% of the N-glycans were biantennary galactosylated (indicated by AA, **Figure 8**). Further, our data indicate that the degree of bigalactosylated peptides is higher in plants with lower expression level of the FTGT (**Figures 7, 8**). On glycopeptides harboring galactoses mass shifts of 132.0423 Da or 264.0846 Da were observed in up to 50% of all detected galactosylated N-glycans. This suggests the linkage of one or two pentoses to the newly introduced β 1,4-linked galactoses (**Figure 8** and **Supplementary Figure 11**). Non-galactosylated glycopeptides carried almost exclusively the complex-type GnGn glycoform. N-glycans bearing plant specific xylose, α 1,3-linked fucose, Le^a epitopes or the corresponding aglycons were not detected for any of the analyzed glycopeptides. Despite the fact that with the introduction of a β 1,4-galactosyltransferase



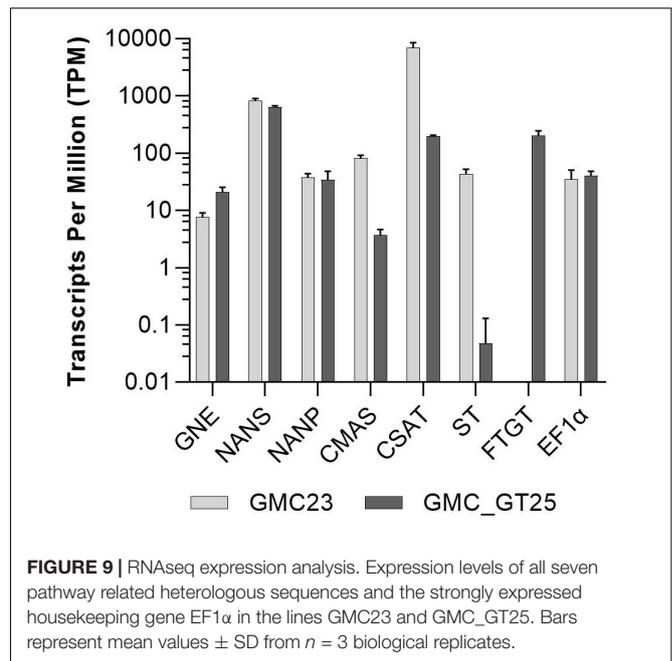
into a CMP-Neu5Ac-producing and CSAT and ST-expressing moss line all requirements for sialylation are addressed, in none of the MS-analyzed lines sialylation of N-glycans was detectable. A corresponding result was reported for *N. benthamiana* by Kittur et al. (2020).

RNAseq-Analysis Revealed the Drop of ST Expression After FTGT Introduction

To address the question why no sialic acid was attached to N-glycans in the GMC_GT lines harboring all genes necessary for sialylation, we analyzed the changes in gene expression in the line GMC_GT25, and its parental line GMC23, without the chimeric β 1,4-galactosyltransferase. RNAseq analysis revealed that the introduction of FTGT led to a decrease of CMAS and CSAT expression of approximately 95% and nearly abolished the ST expression, whereas the expression levels of the first three pathway CDSs remained stable (Figure 9). Even if the reason for the striking drop in ST expression level could not be explained, we assume that this lack of expression is responsible for the absence of sialylation in the GMC_GT lines.

Lines Strongly Expressing the FTST Are Able to Perform Stable N-Glycan Sialylation

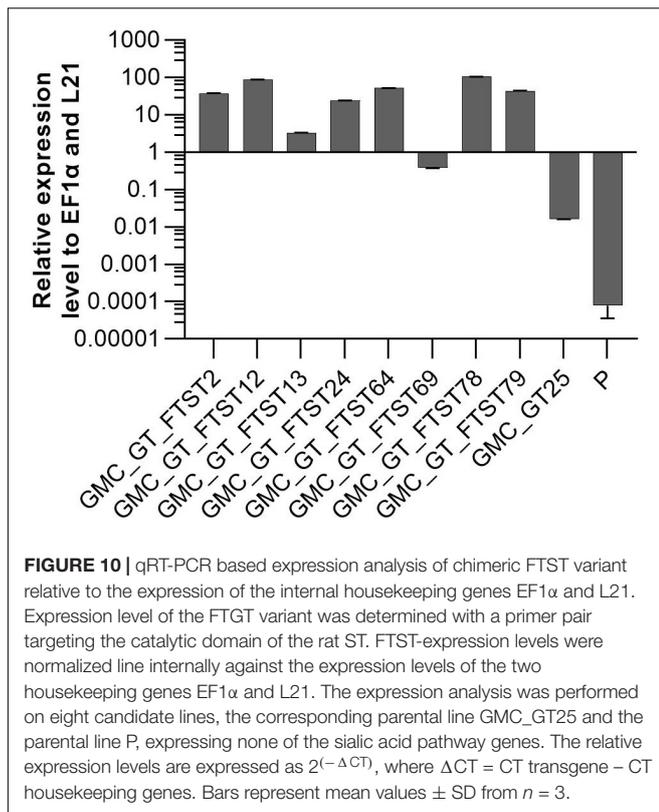
To overcome the absence of ST expression in GMC_GT25, the line was transformed with a chimeric ST version. In this variant, the original CTS of the rat ST was replaced by the



CTS of the endogenous α 1,4-fucosyltransferase already used for the FTGT, giving rise to the chimeric variant FTST. The integration of the FTST was targeted to the P4H2 locus, which is part of the plant-prolylhydroxylases family (Parsons et al., 2013). Plants surviving the hygromycin-based selection were analyzed for targeted construct integration within the P4H2 locus, which could be confirmed for eight GMC_GT_FTST lines (GMC_GT_FTST2, 12, 13, 24, 64, 69, 78, and 79; Supplementary Figure 12). Although these plants express several transgenes and display multiple modifications, no obvious impairment in growth could be assessed for protonema suspension cultures under standard cultivation conditions.

FTST expression was quantified via qRT-PCR in all GMC_GT_FTST lines with targeted genome integration. The used primers were targeted to the catalytic domain of the sialyltransferase, thus amplifying both ST variants. Therefore, the lines GMC_GT_FTST, carrying both the native and the chimeric ST were compared to both, their parental line GMC_GT25, poorly expressing the native ST version and the parental line P, expressing none of the sialic acid pathway genes. As previously detected, the GMC_GT25 displayed only a very weak ST expression, while six out of the eight tested GMC_GT_FTST lines strongly express the chimeric sialyltransferase, with higher expression levels than those of the strongly expressed housekeeping genes EF1 α and L21, while in P no ST-expression could be detected (Figure 10).

To investigate FTST enzyme activity, three lines with FTST expression levels ranging from low to high (GMC_GT_FTST13, 64 and 78) were chosen for MS-based glycopeptide analysis. This analysis revealed that the galactosylation efficiency increased from 60% in the GMC_GT25 parental line to 85–89% in FTST expressing lines (Supplementary Figure 13). Moreover, a higher share of biantennary galactosylated N-glycans could be detected



in all analyzed lines after the introduction of the FTST, increasing from 6.3% in GMC23_GT25 to, e.g., 33.7% in GMC_GT_FTST64 (Supplementary Figure 13). Further, in the two moderate and strong FTST expressing lines GMC_GT_FTST64 and 78 stable *N*-glycan sialylation could be confirmed, whereas in the weak-expressing line GMC_GT_FTST13 no *N*-glycan sialylation was detectable (Figure 11A and Supplementary Figure 13). This validates the activity of the introduced FTST variant. Verification of *N*-glycan sialylation was performed via the detection of defined glycopeptide *m/z*-ratios on MS¹ level (Supplementary Figures 14–17 and Supplementary Table 2) as well as the detection of sialic acid reporter ions [Neu5Ac]⁺ with *m/z* 292.103 and [Neu5Ac - H₂O]⁺ with *m/z* 274.092 and fragments of the corresponding peptide backbone on MS² level (Figures 11B–D and Supplementary Figures 14–17). In GMC_GT_FTST78, the line with the highest FTST expression level, sialylation could be detected in all three analyzed glycopeptides (Figures 11B–D and Supplementary Figure 18). Altogether, sialic acid was linked to 6.3% of all detected glycopeptides, of which 6.0% were NaM, while the remaining 0.3% displayed NaGn structures. Further, 70.8% AM and 0.8% AGn structures were detected, but almost half of them were decorated with one or two pentoses. Interestingly, no pentoses were found when the *N*-glycans were sialylated. Biantennary galactosylation was present in 11.3% of the glycopeptides, but no sialylation of this structure could be detected. The glycosylation patterns for each analyzed *N*-glycosylation site are depicted in Supplementary Figure 18. GMC_GT_FTST64, the analyzed line

with the intermediate FTST-expression level, displayed an overall sialylation efficiency of 1% NaM structures (Supplementary Figure 13), indicating that the efficiency of sialylation is dependent on the expression level of the sialyltransferase. These results verified the correct synthesis, localization and sequential activity of seven heterologous mammalian enzymes and confirmed that stable *N*-glycan sialylation is possible in moss.

DISCUSSION

Since many therapeutics are sialylated glycoproteins, the ability to perform sialylation on complex-type *N*-glycans is highly required for biopharmaceutical production platforms. *N*-glycan sialylation plays an important role for the plasma half-life of glycoproteins, as it protects them from clearance (reviewed in Varki, 2008), which is highly favorable to achieve long lasting therapeutic effects.

To achieve *N*-glycan sialylation in *Physcomitrella* several challenges need to be addressed, namely the production of free sialic acid (Neu5Ac) from the precursor UDP-*N*-acetylglucosamine, which is also present in plants (Stanley et al., 2017), its activation to CMP-Neu5Ac, its transport to the Golgi apparatus and finally its transfer onto a β 1,4-galactosylated *N*-glycan. This requires the coordinated activity of seven mammalian enzymes localized in three different plant subcellular compartments including nucleus, cytosol and Golgi apparatus. Stable integration of the respective genes was directed via gene targeting, which is a frequently used tool for genome engineering in *Physcomitrella* (e.g., Khraiweh et al., 2010; Sakakibara et al., 2013; Decker et al., 2015; Horst et al., 2016; Wiedemann et al., 2018). These knock-in approaches were aimed for selection purposes or to avoid undesired plant-typical posttranslational modifications. The GNN construct was targeted to the APT locus, encoding adenine phosphoribosyltransferase. This enzyme is involved in the purine salvage pathway and recycles adenine into AMP, but can also recognize adenine analogs as substrates, which are toxic when metabolized (Schaff, 1994). Knock-out of the single gene coding for this enzyme in *Physcomitrella* results in resistance to adenine analogs as 2-fluoro adenine; an effect that has been employed as a marker for homologous recombination frequency (Trouiller et al., 2007; Kamisugi et al., 2012). We used this strategy to positively select GNN-transformed plants avoiding the introduction of an additional foreign gene for selection. The integration of the CCSB- and FTST-expression cassettes was targeted to the genes coding for the prolyl-4-hydroxylase 1 (P4H1) and the prolyl-4-hydroxylase 2 (P4H2), respectively, which are members of the plant P4H-family, responsible for hydroxylation of prolines in consensus sequences different to those recognized in humans (Parsons et al., 2013). Moreover, as hydroxyproline serves as anchor for plant-typical *O*-glycosylation, the formation of these structures should be avoided in biopharmaceuticals. Besides, the generation of the potential immunogenic Le^a-epitope was abolished in *Physcomitrella* by targeting the FTGT-expression cassette to the gene encoding β 1,3-galactosyltransferase 1. The knock-out of this gene lead to rEPO devoid of Le^a structures

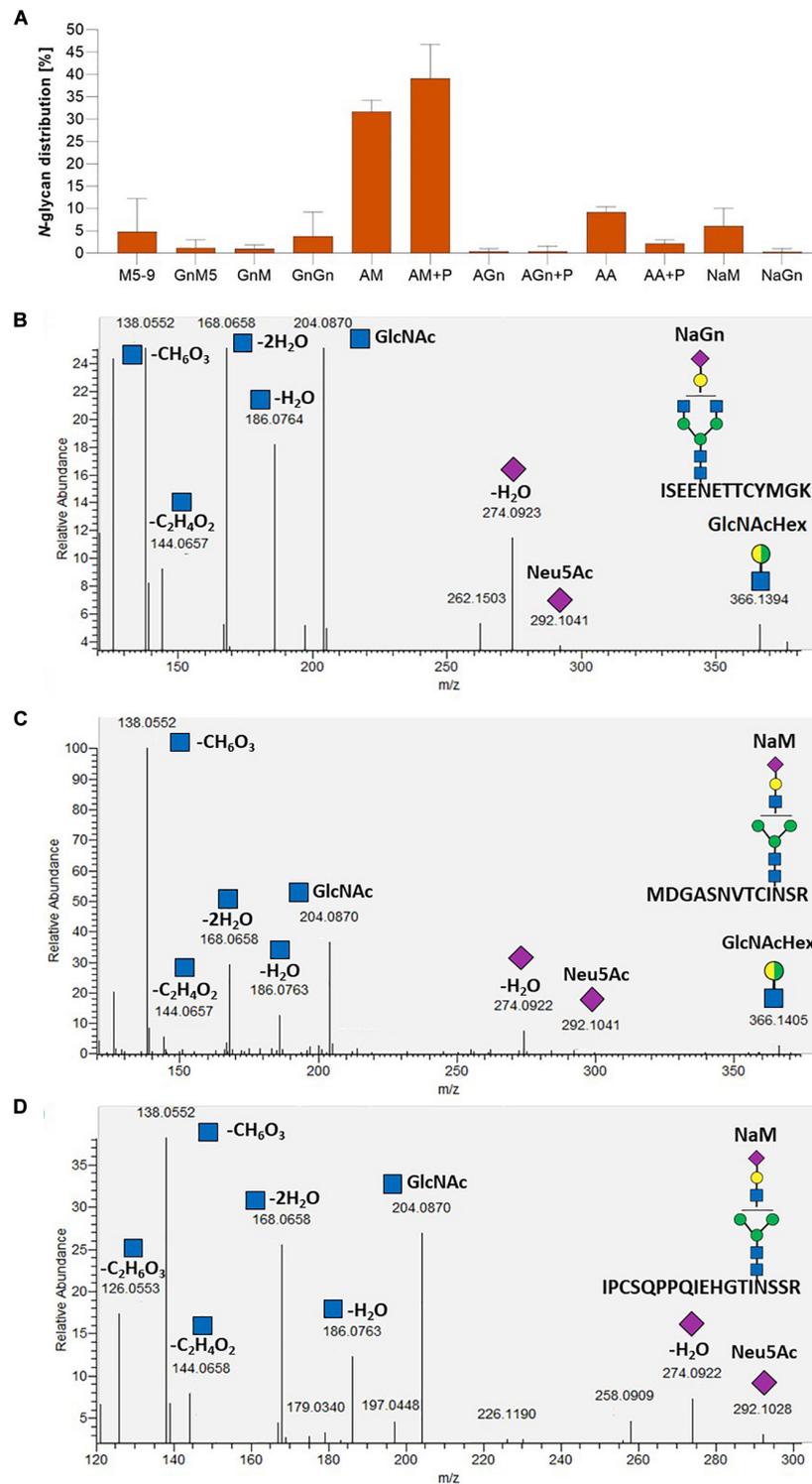


FIGURE 11 | Mass spectrometric analysis of the glycosylation pattern in GMC_GT_FTST78 and MS²-based verification of sialylated glycopeptides. **(A)** MS-based relative quantification of glycopeptides identified on the tryptic digested reporter glycoprotein of the complete sialylation line GMC_GT_FTST78 with a chimeric sialyltransferase. Relative quantification was based on peak area integration of extracted ion chromatograms (EICs) on MS¹ level, for which glycopeptide identities were confirmed on MS² level. For quantification, areas of all confirmed peaks per measurement were summed up and the relative percentages are given for each identified glycan structure. The error bars indicate the standard deviation between the three analyzed glycopeptides. **(B–D)** MS²-based verification of *N*-glycan sialylation on all three analyzed tryptic glycopeptides by the identification of *N*-acetylglucosamine (GlcNAc) and sialic acid (Neu5Ac) reporter ions with the following (Continued)

FIGURE 11 | Continued

m/z-values: [GlcNAc]⁺ = 204.087, [GlcNAc - H₂O]⁺ = 186.076, [GlcNAc - 2H₂O]⁺ = 168.066, [GlcNAc - C₂H₄O₂]⁺ = 144.065, [GlcNAc - CH₆O₃]⁺ = 138.055, [GlcNAc - C₂H₆O₃]⁺ = 126.055, [Neu5Ac]⁺ = 292.103, [Neu5Ac - H₂O]⁺ = 274.092 and the detection of the glycan fragment ion [GlcNAcHex]⁺ = 366.139. M, mannose (green circle); Gn, *N*-acetylglucosamine (blue square); A, galactose (yellow circle); Na, sialic acid (purple rhombos); Hex, hexose (yellow and green circle); stands for the presence of either mannose or galactose).

without any obvious impact on the plant's growth under production conditions (Parsons et al., 2012).

The production of free sialic acid was achieved in *Physcomitrella* by the simultaneous stable expression of the sequences encoding the first three cytosolic active enzymes of the sialic acid biosynthesis pathway: murine GNE, human NANS and human NANP, resulting in the production of Neu5Ac in 10 times higher amounts than previously reported for plants (Castilho et al., 2008). In *A. thaliana* and *N. benthamiana*, either stably or transiently transformed with GNE and NANS, endogenous activity of a putative plant NANP was sufficient for the synthesis of Neu5Ac (Castilho et al., 2008, 2010). However, our data indicate that in moss additional NANP expression led to higher Neu5Ac production.

As the full-length CMAS-CDS could not be amplified from human cDNA, a shorter version, where 120 bp at the 5' end are missing (Castilho et al., 2008), was cloned and introduced in sialic acid-producing lines. Moss plants expressing the truncated version of CMAS were able to activate sialic acid, confirming its activity in *Physcomitrella*.

The activity of GNE, the first enzyme in the biosynthesis of sialic acid, is known to be inhibited by the CMAS product CMP-Neu5Ac (Kornfeld et al., 1964). Mutated versions of this enzyme (R263L and/or R266Q/R266W) were already used in *N. benthamiana*, insect and CHO cells to overcome this negative feedback and increase sialylation of recombinant proteins (Viswanathan et al., 2005; Bork et al., 2007; Son et al., 2011; Kallolimath et al., 2016). We expressed in moss the GNE_{mut} version carrying the double mutation R263L, R266Q which was previously described to be more active *in vitro* than the single mutated versions (Son et al., 2011). Sialic acid-synthesizing moss plants carrying GNE_{mut} produced similar amounts of sialic acid as plants expressing the native GNE-version, confirming that the R263L, R266Q mutations in the allosteric site of GNE do not affect the epimerase/kinase activities of the enzyme. In contrast to the native GNE-expressing lines, no decrease in the production of sialic acid was observed in GNE_{mut}-expressing lines after the introduction of CMAS, revealing that the negative feed-back loop of GNE was successfully prevented in moss. Consequently, up to 25 times more CMP-Neu5Ac was achieved in moss plants expressing the GNE_{mut} compared to lines with the native GNE.

The 2A peptide sequences derived from the *Picornaviridae* family are known to cause a ribosomal skipping event during the process of translation (Szymczak et al., 2004), resulting in isolated proteins from a single open reading frame. The ability of *Physcomitrella* to recognize the 2A peptide sequence from the foot-and-mouth disease virus has already been reported (Pan et al., 2015). We used the P2A sequence (Kim et al., 2011), originated from the porcine teschovirus-1, another virus belonging to this family, to separate the transgene of interest, in our case the chimeric β 1,4-galactosyltransferase, from the *ble*

resistance gene. This resulted in the synthesis of independent FTGT and the Bleomycin-resistance protein, verified by their corresponding enzymatic activities.

Different qualities and degrees of galactosylation have been reported so far after the introduction of a β 1,4-galactosyltransferase in plants. Enzymes involved in *N*-glycosylation are localized in the ER and/or in the Golgi apparatus according to their sequential manner of action in the *N*-glycan maturation, and this localization is determined by their N-terminal CTS region (Saint-Jore-Dupas et al., 2006). Early incorporation of galactose into the maturing *N*-glycan before the action of mannosidase II leads to hybrid-type glycans (Schachter et al., 1983; Hesselink et al., 2014). Therefore, we aimed to target the β 1,4-galactosyltransferase activity to the *trans*-Golgi apparatus by creating a chimeric enzyme consisting of the catalytic domain of the human β 1,4-galactosyltransferase and the localization-determining CTS region of the moss endogenous α 1,4-fucosyltransferase (FT4), which is the last enzyme known to act on the plant *N*-glycan maturation (Saint-Jore-Dupas et al., 2006). Lines expressing this synthetic GalT4 variant (FTGT) displayed up to 60% of galactosylation on the reporter glycopeptides. Mono- and biantennary galactosylated glycans were detected, up to 50% and 15%, respectively. In addition, in some of the galactosylated *N*-glycans the appearance of mass shifts corresponding to the molecular weight of one or two pentoses occurred, indicating that β 1,4-galactosylated *N*-glycans were decorated with additional sugars of thus far unknown identity. The characterization of these residues remains a future task. A very similar galactosylation pattern, including the appearance of pentoses, was described by Kittur et al. (2020) on rEPO in transgenic *N. tabacum* lines expressing a chimeric GalT4 with the CTS region of the rat α 2,6-sialyltransferase (STGT). Interestingly, in transgenic Δ xt/ft *N. benthamiana* lines the same STGT variant led to predominantly bigalactosylated *N*-glycan structures on a transiently co-expressed antibody (Strasser et al., 2009). However, the extent of biantennary galactosylation also seems to depend on the investigated glycoprotein, as in contrast to the high galactosylation efficiencies obtained on antibodies, poor levels were achieved for other proteins in the same system (Kriechbaum et al., 2020). In this context, glycosylation might be influenced by the conformation of the target protein and the accessibility of the maturing glycan to the glycosyltransferases in the Golgi apparatus. Further, our results indicate that the expression level of the enzyme plays a role in the quality of galactosylation. Earlier studies performed in *N. benthamiana* suggested that there is an optimal GalT expression level and levels beyond it resulted in higher amounts of immature *N*-glycans (Kallolimath et al., 2018). This is in agreement with our results, since the amount of bigalactosylated *N*-glycans was higher in lines with a lower FTGT expression level. Therefore, an improvement of the glycosylation pattern toward a higher

share of bigalactosylated *N*-glycan structures via modulation of promoter strength or number of inserted copies in stably transformed moss plants is a future task.

Achieving stable sialylation in *Physcomitrella* requires not just the correct subcellular localization and coordinated activity of seven mammalian enzymes involved in this process, but also their optimal expression levels to produce fully processed glycans with terminal sialic acid. Since the same promoter was used to express all of the transgenes, the differences in the expression levels observed between moss transgenic lines should mainly be determined by the number of constructs integrated into the genome or the loci of integration, respectively. The use of homologous flanks to target the integration in *Physcomitrella* do not always limit this recombination process to the intended locus alone. In contrast to transient systems, the process of plant selection according to the expression level of the gene of interest is a necessary step. According to RNAseq data, the introduction of the FTGT resulted additionally in a drop of CMAS and CSAT and a dramatically decrease of ST expression levels. It has been reported that during moss transfection rearrangements may appear, such as small deletions, insertions or concatenation of the constructs, which are characteristic of non-homologous end-joining or homologous recombination pathways (Kamisugi et al., 2006; Murén et al., 2009). These events might be an explanation for the marked decrease in the expression level of these genes, since some copies of the transgenes might be damaged by recombination events. However, we currently cannot explain why ST expression was especially affected.

To circumvent the loss of ST expression, we created and inserted the chimeric gene for FTST into the genome of the production line. The efficient expression of this chimeric ST variant led to stable sialylation of the recombinant human reporter glycoprotein. The observed structures were NaM and NaGn *N*-glycans, or their isomers, appearing on all analyzed glycopeptides. Furthermore, FTST expression led to an increased galactosylation efficiency from 60 up to 89% combined with an increase in the share of biantennary galactosylated *N*-glycans. Although the sialylation levels achieved in this study are still low (around 6.3%), this is an important step toward the production of sialylated, high-quality biopharmaceuticals in *Physcomitrella*. Different strategies need to be addressed to further optimize the expression levels of the genes involved in the sialylation pathway. It is also well known that culture conditions affect the glycosylation pattern in CHO cells, such as culture pH, temperature, media and supplements, operation mode of the bioreactor and feeding strategies (Hossler et al., 2009; Lewis et al., 2016; Wang et al., 2018; Ehret et al., 2019). Therefore, evaluating the influence of cell culture factors might be a further strategy to increase sialylation levels in *Physcomitrella*.

We established stable *N*-glycan sialylation in *Physcomitrella* by the introduction of seven mammalian enzymes, some of them as chimeric versions, resulting in a sialylated co-expressed human glycoprotein. This demonstrates the coordinated activity and correct subcellular localization of the corresponding enzymes within the cytosol, the nucleus and the Golgi apparatus. Stable expression of all transgenes involved in both the production of the reporter protein as well as the glycosylation, enables the

complete characterization of the production line and ensures consistent product quality.

The now established ability of stably transformed *Physcomitrella* to sialylate proteins widens the range of *N*-glycosylation patterns attainable in this system considerably. This facilitates the customized production of recombinant biopharmaceuticals with various degrees of post-translational modifications to ensure maximal activity. This makes moss a versatile and attractive production system within the biopharmaceutical industry.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA665456.

AUTHOR CONTRIBUTIONS

LB performed most of the experiments. SH performed the MS data analysis. CR performed the RNAseq data analysis. TL cloned the GNN, GM, and CCSB constructs. FRJ cloned the FTGT construct. RF performed the mass-spectrometric based Neu5Ac and CMP-Neu5Ac analyses. LB, JP, NR-M, FA, BF, RR, and ED designed the study and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Excellence Initiative and Strategy of the German Federal and State Governments (GSC-4 SGBM to FRJ and EXC 294 BIOS to RR), Eleva GmbH (to RR) and the Deutscher Akademischer Austauschdienst DAAD (to NR-M).

ACKNOWLEDGMENTS

We thank Agnes Novakovic, Ingrid Heger, Kai Clemens, Carsten Gaa and Maxime Karrer for their support to this work, Prof. Dr. Mitsuyasu Hasebe for kindly providing the BSD plasmid, Dr. Karsten Häffner for providing human and rat cDNA, Dr. Asifa Akhtar for providing murine cDNA, and the Genomics Unit of Instituto Gulbenkian de Ciencia for RNAseq analyses. We also thank Prof. Dr. Bettina Warscheid for the possibility to use the QExactive Plus instrument and Anne Katrin Prowse for proof-reading of the article. CR received funding by the Ministry of Science, Research and Art of the Federal State of Baden-Württemberg (MWK) as part of the Science Data Center funding program BioDATEN.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.610032/full#supplementary-material>

REFERENCES

- Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. doi: 10.1093/nar/gky379
- Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available online at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed August 30, 2020).
- Aoki, K., Ishida, N., and Kawakita, M. (2003). Substrate recognition by nucleotide sugar transporters further characterization of substrate recognition regions by analyses of UDP-galactose/CMP-sialic acid transporter chimeras and biochemical analysis of the substrate specificity of parental and chimeric transporters. *J. Biol. Chem.* 278, 22887–22893. doi: 10.1074/jbc.M302620200
- Bakker, H., Bardor, M., Molthoff, J. W., Gomord, V., Elbers, I., Stevens, L. H., et al. (2001). Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. U.S.A.* 98, 2899–2904. doi: 10.1073/pnas.031419998
- Bakker, H., Routier, F., Ashikov, A., Neumann, D., Bosch, D., and Gerardy-Schahn, R. (2008). A CMP-sialic acid transporter cloned from *Arabidopsis thaliana*. *Carbohydrate Res.* 343, 2148–2152. doi: 10.1016/j.carres.2008.01.010
- Bakker, H., Rouwendal, G. J. A., Karnoup, A. S., Florack, D. E. A., Stoop, G. M., Helsper, J. P. F. G., et al. (2006). An antibody produced in tobacco expressing a hybrid β -1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7577–7582. doi: 10.1073/pnas.0600879103
- Bardor, M., Faveeuw, C., Fitchette, A.-C., Gilbert, D., Galas, L., Trottein, F., et al. (2003). Immunoreactivity in mammals of two typical plant glyco-epitopes, core α (1,3)-fucose and core xylose. *Glycobiology* 13, 427–434. doi: 10.1093/glycob/cwg024
- Baym, M., Kryazhimskiy, S., Lieberman, T. D., Chung, H., Desai, M. M., and Kishony, R. (2015). Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* 10:e0128036. doi: 10.1371/journal.pone.0128036
- Beike, A. K., Lang, D., Zimmer, A. D., Wüst, F., Trautmann, D., Wiedemann, G., et al. (2015). Insights from the cold transcriptome of *Physcomitrella patens*: global specialization pattern of conserved transcriptional regulators and identification of orphan genes involved in cold acclimation. *New Phytol.* 205, 869–881. doi: 10.1111/nph.13004
- Bork, K., Reutter, W., Weidemann, W., and Horstkorte, R. (2007). Enhanced sialylation of EPO by overexpression of UDP-GlcNAc 2-epimerase/ManAc kinase containing a sialuria mutation in CHO cells. *FEBS Lett.* 581, 4195–4198. doi: 10.1016/j.febslet.2007.07.060
- Bosch, D., Castilho, A., Loos, A., Schots, A., and Steinkellner, H. (2013). N-glycosylation of plant-produced recombinant proteins. *Curr. Pharm. Des.* 19, 5503–5512. doi: 10.2174/1381612811319310006
- Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34, 525–527. doi: 10.1038/nbt.3519
- Castilho, A., Gatteringer, P., Grass, J., Jez, J., Pabst, M., Altmann, F., et al. (2011). N-glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans. *Glycobiology* 21, 813–823. doi: 10.1093/glycob/cwr009
- Castilho, A., Pabst, M., Leonard, R., Veit, C., Altmann, F., Mach, L., et al. (2008). Construction of a functional CMP-sialic acid biosynthesis pathway in *Arabidopsis*. *Plant Physiol.* 147, 331–339. doi: 10.1104/pp.108.117572
- Castilho, A., Strasser, R., Stadlmann, J., Grass, J., Jez, J., Gatteringer, P., et al. (2010). In planta protein sialylation through overexpression of the respective mammalian pathway. *J. Biol. Chem.* 285, 15923–15930. doi: 10.1074/jbc.M109.088401
- Chou, H.-H., Hayakawa, T., Diaz, S., Krings, M., Indriati, E., Leakey, M., et al. (2002). Inactivation of CMP-N-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11736–11741. doi: 10.1073/pnas.182257399
- Chung, C. H., Mirakhur, B., Chan, E., Le, Q.-T., Berlin, J., Morse, M., et al. (2008). Cetuximab-induced anaphylaxis and IgE specific for galactose- α -1,3-galactose. *N. Engl. J. Med.* 358, 1109–1117. doi: 10.1056/NEJMoa074943
- Commandeur, U., Twyman, R. M., and Fischer, R. (2003). The biosafety of molecular farming in plants. *AgBiotechNet. Online Serv. Agric. Biotechnol.* 5:9.
- Cox, K. M., Sterling, J. D., Regan, J. T., Gasdaska, J. R., Frantz, K. K., Peele, C. G., et al. (2006). Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol.* 24, 1591–1597. doi: 10.1038/nbt1260
- Decker, E. L., Parsons, J., and Reski, R. (2014). Glyco-engineering for biopharmaceutical production in moss bioreactors. *Front. Plant Sci.* 5:346. doi: 10.3389/fpls.2014.00346
- Decker, E. L., and Reski, R. (2012). Glycoprotein production in moss bioreactors. *Plant Cell Rep.* 31, 453–460. doi: 10.1007/s00299-011-1152-5
- Decker, E. L., and Reski, R. (2020). Mosses in biotechnology. *Curr. Opin. Biotechnol.* 61, 21–27. doi: 10.1016/j.copbio.2019.09.021
- Decker, E. L., Wiedemann, G., and Reski, R. (2015). “Gene targeting for Precision glyco-engineering: production of biopharmaceuticals devoid of plant-typical glycosylation in moss bioreactors,” in *Glyco-Engineering: Methods and Protocols Methods in Molecular Biology*, ed. A. Castilho (New York, NY: Springer), 213–224. doi: 10.1007/978-1-4939-2760-9_15
- Donnelly, M. L. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D., et al. (2001). Analysis of the aphthovirus 2A/2B polyprotein ‘cleavage’ mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal ‘skip’. *J. Gen. Virol.* 82, 1013–1025. doi: 10.1099/0022-1317-82-5-1013
- Ehret, J., Zimmermann, M., Eichhorn, T., and Zimmer, A. (2019). Impact of cell culture media additives on IgG glycosylation produced in Chinese hamster ovary cells. *Biotechnol. Bioeng.* 116, 816–830. doi: 10.1002/bit.26904
- Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354
- Fitchette, A.-C., Cabanes-Macheteau, M., Marvin, L., Martin, B., Satiat-Jeunemaitre, B., Gomord, V., et al. (1999). Biosynthesis and immunolocalization of Lewis a-containing N-glycans in the plant cell. *Plant Physiol.* 121, 333–344. doi: 10.1104/pp.121.2.333
- Frank, W., Decker, E. L., and Reski, R. (2005). Molecular tools to study *Physcomitrella patens*. *Plant Biol.* 7, 220–227. doi: 10.1055/s-2005-865645
- Fujiyama, K., Furukawa, A., Katsura, A., Masaki, R., Omasa, T., and Seki, T. (2007). Production of mouse monoclonal antibody with galactose-extended sugar chain by suspension cultured tobacco BY2 cells expressing human β (1,4)-galactosyltransferase. *Biochem. Biophys. Res. Commun.* 358, 85–91. doi: 10.1016/j.bbrc.2007.04.054
- Ghaderi, D., Zhang, M., Hurtado-Ziola, N., and Varki, A. (2012). Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. *Biotechnol. Genet. Eng. Rev.* 28, 147–175. doi: 10.5661/bger-28-147
- Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318
- Gomord, V., Chamberlain, P., Jefferis, R., and Faye, L. (2005). Biopharmaceutical production in plants: problems, solutions and opportunities. *Trends Biotechnol.* 23, 559–565. doi: 10.1016/j.tibtech.2005.09.003
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944
- Häffner, K., Parsons, J., Bohlender, L. L., Hoernstein, S., Niederkrüger, H., Fode, B., et al. (2017). Treatment of experimental C3 Glomerulopathy by human complement factor H produced in glycosylation-optimized *Physcomitrella patens*. *Mol. Immunol.* 89:120. doi: 10.1016/j.molimm.2017.06.050
- Halim, A., Westerlind, U., Pett, C., Schorlemer, M., Rüetschi, U., Brinkmalm, G., et al. (2014). Assignment of saccharide identities through analysis of oxonium ion fragmentation profiles in LC–MS/MS of glycopeptides. *J. Proteome Res.* 13, 6024–6032. doi: 10.1021/pr500898r
- Hamadeh, R. M., Jarvis, G. A., Galili, U., Mandrell, R. E., Zhou, P., and Griffiss, J. M. (1992). Human natural anti-Gal IgG regulates alternative complement pathway activation on bacterial surfaces. *J. Clin. Invest.* 89, 1223–1235. doi: 10.1172/JCI115706
- Hanania, U., Ariel, T., Tekoah, Y., Fux, L., Sheva, M., Gubbay, Y., et al. (2017). Establishment of a tobacco BY2 cell line devoid of plant-specific xylose and fucose as a platform for the production of biotherapeutic proteins. *Plant Biotechnol. J.* 15, 1120–1129. doi: 10.1111/pbi.12702

- Hara, S., Takemori, Y., Yamaguchi, M., Nakamura, M., and Ohkura, Y. (1987). Fluorometric high-performance liquid chromatography of *N*-acetyl- and *N*-glycolylneuraminic acids and its application to their microdetermination in human and animal sera, glycoproteins, and glycolipids. *Anal. Biochem.* 164, 138–145. doi: 10.1016/0003-2697(87)90377-0
- Hennermann, J. B., Arash-Kaps, L., Fekete, G., Schaaf, A., Busch, A., and Frischmuth, T. (2019). Pharmacokinetics, pharmacodynamics, and safety of moss- α -Galactosidase A in patients with Fabry disease. *J. Inherit. Metab. Dis.* 42, 527–533. doi: 10.1002/jimd.12052
- Hesselink, T., Rouwendal, G. J. A., Henquet, M. G. L., Florack, D. E. A., Helsper, J. P. F. G., and Bosch, D. (2014). Expression of natural human β 1,4-GalT1 variants and of non-mammalian homologues in plants leads to differences in galactosylation of *N*-glycans. *Transgenic Res.* 23, 717–728. doi: 10.1007/s11248-014-9806-z
- Hintze, S., Limmer, S., Dabrowska-Schlepp, P., Berg, B., Kriehoff, N., Busch, A., et al. (2020). Moss-derived human recombinant GAA provides an optimized enzyme uptake in differentiated human muscle cells of pompe disease. *Int. J. Mol. Sci.* 21:2642. doi: 10.3390/ijms21072642
- Horst, N. A., Katz, A., Pereman, I., Decker, E. L., Ohad, N., and Reski, R. (2016). A single homeobox gene triggers phase transition, embryogenesis and asexual reproduction. *Nature Plants* 2:15209. doi: 10.1038/nplants.2015.209
- Horstmann, V., Huether, C. M., Jost, W., Reski, R., and Decker, E. L. (2004). Quantitative promoter analysis in *Physcomitrella patens*: a set of plant vectors activating gene expression within three orders of magnitude. *BMC Biotechnol.* 4:13. doi: 10.1186/1472-6750-4-13
- Hossler, P., Khattak, S. F., and Li, Z. J. (2009). Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* 19, 936–949. doi: 10.1093/glycob/cwp079
- Huether, C. M., Lienhart, O., Baur, A., Stemmer, C., Gorr, G., Reski, R., et al. (2005). Glyco-engineering of moss lacking plant-specific sugar residues. *Plant Biol. (Stuttg.)* 7, 292–299. doi: 10.1055/s-2005-837653
- Jansing, J., Sack, M., Augustine, S. M., Fischer, R., and Bortesi, L. (2018). CRISPR/Cas9-mediated knockout of six glycosyltransferase genes in *Nicotiana benthamiana* for the production of recombinant proteins lacking β -1,2-xylose and core α -1,3-fucose. *Plant Biotechnol. J.* 17, 350–361. doi: 10.1111/pbi.12981
- Jefferis, R. (2005). glycosylation of recombinant antibody therapeutics. *Biotechnol. Prog.* 21, 11–16. doi: 10.1021/bp040016j
- Jenkins, N., Murphy, L., and Tyther, R. (2008). Post-translational modifications of recombinant proteins: significance for biopharmaceuticals. *Mol. Biotechnol.* 39, 113–118. doi: 10.1007/s12033-008-9049-4
- Jourdian, G. W., Dean, L., and Roseman, S. (1971). The sialic acids XI. A periodate-resorcinol method for the quantitative estimation of free sialic acids and their glycosides. *J. Biol. Chem.* 246, 430–435.
- Kallolimath, S., Castilho, A., Strasser, R., Grünwald-Gruber, C., Altmann, F., Strubl, S., et al. (2016). Engineering of complex protein sialylation in plants. *Proc. Natl. Acad. Sci. U.S.A.* 113, 9498–9503. doi: 10.1073/pnas.1604371113
- Kallolimath, S., Gruber, C., Steinkellner, H., and Castilho, A. (2018). Promoter choice impacts the efficiency of plant glyco-engineering. *Biotechnol. J.* 13:1700380. doi: 10.1002/biot.201700380
- Kamisugi, Y., Schaefer, D. G., Kozak, J., Charlot, F., Vrielynck, N., Holá, M., et al. (2012). MRE11 and RAD50, but not NBS1, are essential for gene targeting in the moss *Physcomitrella patens*. *Nucleic Acids Res.* 40, 3496–3510. doi: 10.1093/nar/gkr1272
- Kamisugi, Y., Schlink, K., Rensing, S. A., Schween, G., von Stackelberg, M., Cuming, A. C., et al. (2006). The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration. *Nucleic Acids Res.* 34, 6205–6214. doi: 10.1093/nar/gkl832
- Khraiwesh, B., Arif, M. A., Seumel, G. I., Ossowski, S., Weigel, D., Reski, R., et al. (2010). Transcriptional control of gene expression by MicroRNAs. *Cell* 140, 111–122. doi: 10.1016/j.cell.2009.12.023
- Kim, J. H., Lee, S.-R., Li, L.-H., Park, H.-J., Park, J.-H., Lee, K. Y., et al. (2011). High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* 6:e18556. doi: 10.1371/journal.pone.0018556
- Kittur, F. S., Hung, C.-Y., Zhu, C., Shajahan, A., Azadi, P., Thomas, M. D., et al. (2020). Glycoengineering tobacco plants to stably express recombinant human erythropoietin with different *N*-glycan profiles. *Int. J. Biol. Macromol.* 157, 158–169. doi: 10.1016/j.ijbiomac.2020.04.199
- Kolarich, D., Jensen, P. H., Altmann, F., and Packer, N. H. (2012). Determination of site-specific glycan heterogeneity on glycoproteins. *Nat. Protocols* 7, 1285–1298. doi: 10.1038/nprot.2012.062
- Komatsu, E., Buist, M., Roy, R., Gomes de Oliveira, A. G., Bodnar, E., Salama, A., et al. (2016). Characterization of immunoglobulins through analysis of *N*-glycopeptides by MALDI-TOF MS. *Methods* 104, 170–181. doi: 10.1016/j.ymeth.2016.01.005
- Koprivova, A., Altmann, F., Gorr, G., Kopriva, S., Reski, R., and Decker, E. L. (2003). *N*-glycosylation in the moss *Physcomitrella patens* is organized similarly to that in higher plants. *Plant Biol.* 5, 582–591. doi: 10.1055/s-2003-44721
- Koprivova, A., Stemmer, C., Altmann, F., Hoffmann, A., Kopriva, S., Gorr, G., et al. (2004). Targeted knockouts of *Physcomitrella* lacking plant-specific immunogenic *N*-glycans. *Plant Biotechnol. J.* 2, 517–523. doi: 10.1111/j.1467-7652.2004.00100.x
- Kornfeld, S., Kornfeld, R., Neufeld, E. F., and O'Brien, P. J. (1964). The feedback control of sugar nucleotide biosynthesis in liver. *Proc. Natl. Acad. Sci. U.S.A.* 52, 371–379. doi: 10.1073/pnas.52.2.371
- Kriechbaum, R., Ziaee, E., Grünwald-Gruber, C., Buscaill, P., van der Hoorn, R. A. L., and Castilho, A. (2020). BGAL1 depletion boosts the level of β -galactosylation of *N*- and *O*-glycans in *N. benthamiana*. *Plant Biotechnol. J.* 18, 1537–1549. doi: 10.1111/pbi.13316
- Kubo, M., Imai, A., Nishiyama, T., Ishikawa, M., Sato, Y., Kurata, T., et al. (2013). System for stable β -estradiol-inducible gene expression in the moss *Physcomitrella patens*. *PLoS One* 8:e77356. doi: 10.1371/journal.pone.0077356
- Lang, D., Ullrich, K. K., Murat, F., Fuchs, J., Jenkins, J., Haas, F. B., et al. (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J.* 93, 515–533. doi: 10.1111/tpj.13801
- Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.-C., Gomord, V., and Faye, L. (1998). *N*-Glycoprotein biosynthesis in plants: recent developments and future trends. *Plant Mol. Biol.* 38, 31–48. doi: 10.1023/A:1006012005654
- Lewis, A. M., Croughan, W. D., Aranibar, N., Lee, A. G., Warrack, B., Abu-Absi, N. R., et al. (2016). Understanding and controlling sialylation in a CHO Fc-fusion process. *PLoS One* 11:e0157111. doi: 10.1371/journal.pone.0157111
- Lingg, N., Zhang, P., Song, Z., and Bardor, M. (2012). The sweet tooth of biopharmaceuticals: Importance of recombinant protein glycosylation analysis. *Biotechnol. J.* 7, 1462–1472. doi: 10.1002/biot.201200078
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Macaulay, I. C., Teng, M. J., Haerty, W., Kumar, P., Ponting, C. P., and Voet, T. (2016). Separation and parallel sequencing of the genomes and transcriptomes of single cells using G&T-seq. *Nat. Protoc.* 11, 2081–2103. doi: 10.1038/nprot.2016.138
- Maliekal, P., Vertommen, D., Delpierre, G., and Van Schaftingen, E. (2006). Identification of the sequence encoding *N*-acetylneuraminase-9-phosphate phosphatase. *Glycobiology* 16, 165–172. doi: 10.1093/glycob/cwj050
- Margolin, E., Chapman, R., Williamson, A., Rybicki, E. P., and Meyers, A. E. (2018). Production of complex viral glycoproteins in plants as vaccine immunogens. *Plant Biotechnol. J.* 16, 1531–1545. doi: 10.1111/pbi.12963
- Margolin, E. A., Strasser, R., Chapman, R., Williamson, A.-L., Rybicki, E. P., and Meyers, A. E. (2020). Engineering the plant secretory pathway for the production of next-generation pharmaceuticals. *Trends Biotechnol.* 38, 1034–1044. doi: 10.1016/j.tibtech.2020.03.004
- Merx, S., Smargiasso, N., Chaumont, F., De Pauw, E., Boutry, M., and Navarre, C. (2017). Inactivation of the β (1,2)-xylosyltransferase and the α (1,3)-fucosyltransferase genes in *Nicotiana tabacum* BY-2 Cells by a multiplex CRISPR/Cas9 strategy results in glycoproteins without plant-specific glycans. *Front. Plant Sci.* 8:403. doi: 10.3389/fpls.2017.00403
- Merlin, M., Gecchele, E., Capaldi, S., Pezzotti, M., and Avesani, L. (2014). Comparative evaluation of recombinant protein production in different biofactories: the green perspective. *BioMed. Res. Int.* 2014:136419. doi: 10.1155/2014/136419
- Michelfelder, S., Parsons, J., Bohlender, L. L., Hoernstein, S. N. W., Niederkrüger, H., Busch, A., et al. (2017). Moss-produced, glycosylation-optimized human factor H for therapeutic application in complement disorders. *J. Am. Soc. Nephrol. JASN* 28:1462. doi: 10.1681/ASN.2015070745

- Misaki, R., Kimura, Y., Palacpac, N. Q., Yoshida, S., Fujiyama, K., and Seki, T. (2003). Plant cultured cells expressing human β 1,4-galactosyltransferase secrete glycoproteins with galactose-extended N-linked glycans. *Glycobiology* 13, 199–205. doi: 10.1093/glycob/cwg021
- Montero-Morales, L., and Steinkellner, H. (2018). Advanced Plant-Based Glycan Engineering. *Front. Bioeng. Biotechnol.* 6:81. doi: 10.3389/fbioe.2018.00081
- Moorkens, E., Meuwissen, N., Huys, I., Declerck, P., Vulto, A. G., and Simoens, S. (2017). The market of biopharmaceutical medicines: a snapshot of a diverse industrial landscape. *Front. Pharmacol.* 8:314. doi: 10.3389/fphar.2017.00314
- Murén, E., Nilsson, A., Ulfstedt, M., Johansson, M., and Ronne, H. (2009). Rescue and characterization of episomally replicating DNA from the moss *Physcomitrella*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19444–19449. doi: 10.1073/pnas.0908037106
- Nguyen, T. S., Misaki, R., Ohashi, T., and Fujiyama, K. (2020). Enhancement of sialylation in rIgG in glyco-engineered Chinese hamster ovary cells. *Cytotechnology* 72, 343–355. doi: 10.1007/s10616-020-00381-z
- Paccalet, T., Bardor, M., Rihouey, C., Delmas, F., Chevalier, C., D'Aoust, M.-A., et al. (2007). Engineering of a sialic acid synthesis pathway in transgenic plants by expression of bacterial Neu5Ac-synthetizing enzymes. *Plant Biotechnol. J.* 5, 16–25. doi: 10.1111/j.1467-7652.2006.00211.x
- Palacpac, N. Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T., et al. (1999). Stable expression of human β 1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *Proc. Natl. Acad. Sci. U.S.A.* 96, 4692–4697. doi: 10.1073/pnas.96.8.4692
- Pan, X.-W., Han, L., Zhang, Y.-H., Chen, D.-F., and Simonsen, H. T. (2015). Sclareol production in the moss *Physcomitrella patens* and observations on growth and terpenoid biosynthesis. *Plant Biotechnol. Rep.* 9, 149–159. doi: 10.1007/s11816-015-0353-8
- Parsons, J., Altmann, F., Arrenberg, C. K., Koprivova, A., Beike, A. K., Stemmer, C., et al. (2012). Moss-based production of asialo-erythropoietin devoid of Lewis A and other plant-typical carbohydrate determinants. *Plant Biotechnol. J.* 10, 851–861. doi: 10.1111/j.1467-7652.2012.00704.x
- Parsons, J., Altmann, F., Graf, M., Stadlmann, J., Reski, R., and Decker, E. L. (2013). A gene responsible for prolyl-hydroxylation of moss-produced recombinant human erythropoietin. *Sci. Rep.* 3:3019. doi: 10.1038/srep03019
- Picelli, S., Faridani, O. R., Björklund, Å.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* 9, 171–181. doi: 10.1038/nprot.2014.006
- Reinke, S. O., Lehmer, G., Hinderlich, S., and Reutter, W. (2009). Regulation and pathophysiological implications of UDP-GlcNAc 2-epimerase/ManNAc kinase (GNE) as the key enzyme of sialic acid biosynthesis. *Biol. Chem.* 390, 591–599. doi: 10.1515/BC.2009.073
- Reski, R., Bae, H., and Simonsen, H. T. (2018). *Physcomitrella patens*, a versatile synthetic biology chassis. *Plant Cell Reports* 37, 1409–1417. doi: 10.1007/s00299-018-2293-6
- Roseman, S., Jourdain, G. W., Watson, D., and Rood, R. (1961). Enzymatic synthesis of sialic acid 9-phosphates. *Proc. Natl. Acad. Sci. U.S.A.* 47, 958–961. doi: 10.1073/pnas.47.7.958
- Rozov, S. M., Permyakova, N. V., and Deineko, E. V. (2018). Main strategies of plant expression system glycoengineering for producing humanized recombinant pharmaceutical proteins. *Biochem. Moscow* 83, 215–232. doi: 10.1134/S0006297918030033
- Saint-Jore-Dupas, C., Nebenführ, A., Boulaflous, A., Follet-Gueye, M.-L., Plasson, C., Hawes, C., et al. (2006). Plant N-glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway. *Plant Cell* 18, 3182–3200. doi: 10.1105/tpc.105.03.6400
- Sakakibara, K., Ando, S., Yip, H. K., Tamada, Y., Hiwatashi, Y., Murata, T., et al. (2013). KNOX2 genes regulate the haploid-to-diploid morphological transition in land plants. *Science* 339, 1067–1070. doi: 10.1126/science.1230082
- Schachter, H., Narasimhan, S., Gleeson, P., and Vella, G. (1983). Control of branching during the biosynthesis of asparagine-linked oligosaccharides. *Can. J. Biochem. Cell Biol.* 61, 1049–1066. doi: 10.1139/o83-134
- Schaff, D. A. (1994). The adenine phosphoribosyltransferase (APRT) selectable marker system. *Plant Sci.* 101, 3–9. doi: 10.1016/0168-9452(94)90159-7
- Schiestl, M., Stangler, T., Torella, C., Čepeljnik, T., Toll, H., and Grau, R. (2011). Acceptable changes in quality attributes of glycosylated biopharmaceuticals. *Nat. Biotechnol.* 29, 310–312. doi: 10.1038/nbt.1839
- Schween, G., Hohe, A., Koprivova, A., and Reski, R. (2003). Effects of nutrients, cell density and culture techniques on protoplast regeneration and early protonema development in a moss, *Physcomitrella patens*. *J. Plant Physiol.* 160, 209–212. doi: 10.1078/0176-1617-00855
- Shen, J.-S., Busch, A., Day, T. S., Meng, X.-L., Yu, C. I., Dabrowska-Schlepp, P., et al. (2016). Mannose receptor-mediated delivery of moss-made α -galactosidase A efficiently corrects enzyme deficiency in Fabry mice. *J. Inherit. Metab. Dis.* 39, 293–303. doi: 10.1007/s10545-015-9886-9
- Shin, Y.-J., Chong, Y.-J., Yang, M.-S., and Kwon, T.-H. (2011). Production of recombinant human granulocyte macrophage-colony stimulating factor in rice cell suspension culture with a human-like N-glycan structure. *Plant Biotechnol. J.* 9, 1109–1119. doi: 10.1111/j.1467-7652.2011.00636.x
- Smith, T., O'Kennedy, M. M., Wandrag, D. B. R., Adeyemi, M., and Abolnik, C. (2020). Efficacy of a plant-produced virus-like particle vaccine in chickens challenged with Influenza A H6N2 virus. *Plant Biotechnol. J.* 18, 502–512. doi: 10.1111/pbi.13219
- Sochaj, A. M., Świdarska, K. W., and Otlewski, J. (2015). Current methods for the synthesis of homogeneous antibody–drug conjugates. *Biotechnol. Adv.* 33, 775–784. doi: 10.1016/j.biotechadv.2015.05.001
- Son, Y.-D., Jeong, Y. T., Park, S.-Y., and Kim, J. H. (2011). Enhanced sialylation of recombinant human erythropoietin in Chinese hamster ovary cells by combinatorial engineering of selected genes. *Glycobiology* 21, 1019–1028. doi: 10.1093/glycob/cwr034
- Sourrouille, C., Marquet-Blouin, E., D'Aoust, M.-A., Kiefer-Meyer, M.-C., Seveno, M., Pagny-Salehabadi, S., et al. (2008). Down-regulated expression of plant-specific glycoepitopes in alfalfa. *Plant Biotechnol. J.* 6, 702–721. doi: 10.1111/j.1467-7652.2008.00353.x
- Stanley, P., Taniguchi, N., and Aebi, M. (2017). “N-Glycans,” in *Essentials of Glycobiology*, eds A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, et al. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press). doi: 10.1101/glycobiology.3e.009
- Stoger, E., Fischer, R., Moloney, M., and Ma, J. K.-C. (2014). Plant molecular pharming for the treatment of chronic and infectious diseases. *Annu. Rev. Plant Biol.* 65, 743–768. doi: 10.1146/annurev-arplant-050213-035850
- Strasser, R., Altmann, F., Mach, L., Glössl, J., and Steinkellner, H. (2004). Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking beta1,2-linked xylose and core alpha1,3-linked fucose. *FEBS Lett.* 561, 132–136. doi: 10.1016/S0014-5793(04)00150-4
- Strasser, R., Bondili, J. S., Vavra, U., Schoberer, J., Svoboda, B., Glössl, J., et al. (2007). A unique β 1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell* 19, 2278–2292. doi: 10.1105/tpc.107.052985
- Strasser, R., Castilho, A., Stadlmann, J., Kunert, R., Quendler, H., Gättinger, P., et al. (2009). Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous β 1,4-galactosylated N-glycan profile. *J. Biol. Chem.* 284, 20479–20485. doi: 10.1074/jbc.M109.014126
- Strasser, R., Stadlmann, J., Schähs, M., Stiegler, G., Quendler, H., Mach, L., et al. (2008). Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* 6, 392–402. doi: 10.1111/j.1467-7652.2008.00330.x
- Szymczak, A. L., Workman, C. J., Wang, Y., Vignali, K. M., Dilioglou, S., Vanin, E. F., et al. (2004). Correction of multi-gene deficiency in vivo using a single “self-cleaving” 2A peptide-based retroviral vector. *Nat. Biotechnol.* 22, 589–594. doi: 10.1038/nbt957
- Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., et al. (2003). Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl. Acad. Sci. U.S.A.* 100, 12045–12050. doi: 10.1073/pnas.2131556100
- Tiralongo, J., Ashikov, A., Routier, F., Eckhardt, M., Bakker, H., Gerardy-Schahn, R., et al. (2006). Functional expression of the CMP-sialic acid transporter in *Escherichia coli* and its identification as a simple mobile carrier. *Glycobiology* 16, 73–81. doi: 10.1093/glycob/cwj029
- Top, O., Parsons, J., Bohlender, L. L., Michelfelder, S., Kopp, P., Busch-Steenberg, C., et al. (2019). Recombinant production of MFHR1, a novel synthetic

- multitarget complement inhibitor, in moss bioreactors. *Front. Plant Sci.* 10:260. doi: 10.3389/fpls.2019.00260
- Trouiller, B., Charlot, F., Choinard, S., Schaefer, D. G., and Nogué, F. (2007). Comparison of gene targeting efficiencies in two mosses suggests that it is a conserved feature of Bryophyte transformation. *Biotechnol. Lett.* 29, 1591–1598. doi: 10.1007/s10529-007-9423-5
- Tschofen, M., Knopp, D., Hood, E., and Stöger, E. (2016). Plant molecular farming: much more than medicines. *Annu. Rev. Anal. Chem.* 9, 271–294. doi: 10.1146/annurev-anchem-071015-041706
- Twyman, R. M., Stoger, E., Schillberg, S., Christou, P., and Fischer, R. (2003). Molecular farming in plants: host systems and expression technology. *Trends Biotechnol.* 21, 570–578. doi: 10.1016/j.tibtech.2003.10.002
- van Dussen, L., Zimran, A., Akkerman, E. M., Aerts, J. M. F. G., Petakov, M., Elstein, D., et al. (2013). Taliglucerase alfa leads to favorable bone marrow responses in patients with type I Gaucher disease. *Blood Cells Mol. Dis.* 50, 206–211. doi: 10.1016/j.bcmd.2012.11.001
- van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J. P., Loutelier-Bourhis, C., Rayon, C., et al. (2000). Beta(1,2)-xylose and alpha(1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J. Biol. Chem.* 275, 11451–11458. doi: 10.1074/jbc.275.15.11451
- Varki, A. (2008). Sialic acids in human health and disease. *Trends Mol. Med.* 14, 351–360. doi: 10.1016/j.molmed.2008.06.002
- Varki, A., Schnaar, R. L., and Schauer, R. (2017). “Sialic Acids and other nonulosonic acids,” in *Essentials of Glycobiology*, eds A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, et al. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press). doi: 10.1101/glycobiology.3e.015
- Viëtor, R., Loutelier-Bourhis, C., Fitchette, A.-C., Margerie, P., Gonneau, M., Faye, L., et al. (2003). Protein N-glycosylation is similar in the moss *Physcomitrella patens* and in higher plants. *Planta* 218, 269–275. doi: 10.1007/s00425-003-1099-z
- Viswanathan, K., Narang, S., Hinderlich, S., Lee, Y. C., and Betenbaugh, M. J. (2005). Engineering intracellular CMP-sialic acid metabolism into insect cells and methods to enhance its generation. *Biochemistry* 44, 7526–7534. doi: 10.1021/bi047477y
- Walsh, G. (2018). Biopharmaceutical benchmarks 2018. *Nat. Biotechnol.* 36, 1136–1145. doi: 10.1038/nbt.4305
- Wang, Q., Chung, C.-Y., Chough, S., and Betenbaugh, M. J. (2018). Antibody glycoengineering strategies in mammalian cells. *Biotechnol. Bioeng.* 115, 1378–1393. doi: 10.1002/bit.26567
- Wiedemann, G., Gessel, N., van, Köchl, F., Hunn, L., Schulze, K., Malouk, L., et al. (2018). RecQ helicases function in development, DNA repair, and gene targeting in *Physcomitrella patens*. *Plant Cell* 30, 717–736. doi: 10.1105/tpc.17.00632
- Wilson, I. B. H., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J. M., Kamerling, J. P., et al. (2001). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core α 1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261
- Zeleny, R., Kolarich, D., Strasser, R., and Altmann, F. (2006). Sialic acid concentrations in plants are in the range of inadvertent contamination. *Planta* 224, 222–227. doi: 10.1007/s00425-005-0206-8
- Zhang, Z., Shah, B., and Richardson, J. (2019). Impact of Fc N-glycan sialylation on IgG structure. *mAbs* 11, 1381–1390. doi: 10.1080/19420862.2019.1655377
- Zhang, Z., Wuhrer, M., and Holst, S. (2018). Serum sialylation changes in cancer. *Glycoconj. J.* 35, 139–160. doi: 10.1007/s10719-018-9820-0
- Zhao, W., Chen, T.-L. L., Vertel, B. M., and Colley, K. J. (2006). The CMP-sialic acid transporter is localized in the medial-trans Golgi and possesses two specific endoplasmic reticulum export motifs in its carboxyl-terminal cytoplasmic tail. *J. Biol. Chem.* 281, 31106–31118. doi: 10.1074/jbc.M605564200

Conflict of Interest: The authors are inventors of patents and patent applications related to the production of recombinant proteins in moss. RR is an inventor of the moss bioreactor and a founder of Greenovation Biotech, now Eleva GmbH. He currently serves as advisory board member of this company. BF is an employee of the company Eleva GmbH.

Copyright © 2020 Bohlender, Parsons, Hoernstein, Rempfer, Ruiz-Molina, Lorenz, Rodríguez Jahnke, Figl, Fode, Altmann, Reski and Decker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.