

## **Aufklärung der Genstruktur der Prenyltransferasen aus dem Bittersäurebiosyntheseweg der Hopfenpflanze**

Michael H. Hagemann\*, Theodore Terence Tarudji, Patrick Winterhagen, Anton Lutz, Elisabeth Seigner, Gerd Weber, Jens N. Wünsche

DGG-Proceedings, Vol. 9, 2019, No. 6, p. 1-5.  
DOI: 10.5288/dgg-pr-mh-2019

\*Korrespondierender Autor:

Michael Helmut Hagemann  
Fg. Ertragsphysiologie der Sonderkulturen  
Universität Hohenheim  
Emil-Wolff-Str. 25  
70599 Stuttgart  
Germany

Email: michael@uni-hohenheim.de

## **Aufklärung der Genstruktur der Prenyltransferasen aus dem Bittersäurebiosyntheseweg der Hopfenpflanze**

Michael H. Hagemann<sup>a</sup>, Theodore Terence Tarudji<sup>a</sup>, Patrick Winterhagen<sup>a</sup>, Anton Lutz<sup>b</sup>,  
Elisabeth Seigner<sup>b</sup>, Gerd Weber<sup>a</sup>, Jens N. Wünsche<sup>a</sup>

<sup>a</sup>Fachgebiet Ertragsphysiologie der Sonderkulturen, Universität Hohenheim,  
Emil-Wolff-Straße 25, 70599 Stuttgart, Deutschland

<sup>b</sup>Züchtungsforschung Hopfen, Bayerische Landesanstalt für Landwirtschaft,  
Hüll, 85283 Wolnzach, Deutschland

### **1. Einleitung, Stand des Wissens, Zielsetzung**

Deutschland ist neben den USA der zweitgrößte Produzent von Hopfen für die Brauindustrie. Dabei macht die Gruppe der Bitterhopfen im deutschen Anbau etwa die Hälfte der Erntemenge aus. Die Erntequalität und somit der Preis von Bitterhopfen ist dabei entscheidend von der Konzentration an Bittersäuren in den Hopfendolden abhängig. Die Bittersäurekonzentration hängt hauptsächlich von der Aktivität der drei nur im Hopfen vorkommenden Gene der Bittersäurebiosynthese ab; der Valerophenone Synthase (VPS) (Okada and Ito, 2001), sowie der Prenyltransferase 1 (PT1) und der Prenyltransferase 2 (PT2) (Li et al., 2015; Tsurumaru et al., 2010). Durch die künstliche Nachbildung der Bittersäurebiosynthese in transgenen Hefen konnte schlüssig belegt werden, dass ohne die VPS, PT1 und PT2 die Bittersäurebiosynthese nicht möglich ist (Li et al., 2015; Tsurumaru et al., 2010). Die relativ einfache genomische Struktur des VPS-Gens wurde bereits 2001 aufgeklärt (Okada and Ito, 2001) und es gab bereits erste Ansätze VPS-Sequenzvarianten mit der Bittersäurekonzentration zu korrelieren (Castro et al., 2008). Für die PT1 und PT2 wurden solche Untersuchungen bislang noch nicht durchgeführt, obwohl es sich hierbei um genauso zentrale Elemente der Bittersäurebiosynthese handelt wie bei der VPS. Neben akademischem Interesse an den Genen der Bittersäurebiosynthese sind diese Gene wichtige Ziele für die Züchtung neuer Bitterhopfensorten, da gezeigt wurde, dass der wichtigste Sekundärmetabolit für die Bittere, die Alphasäure, eine sehr hohe Heritabilität von 0.80 bis 0.94 aufweist und somit ein enger Zusammenhang zwischen der Gensequenz und der Konzentration von Alphasäure besteht (Beatson and Alspach, 2009). Generell wird im Hopfen wie bei anderen Kulturpflanzen in der Züchtung zunehmend auf Marker-gestützte Selektion gesetzt, um schneller und kosteneffizienter neue Sorten bereitstellen zu können. Typischerweise wird in der Pflanzenzüchtung von Hauptkulturen hierzu auf die Erfassung von genetischer Sequenzvarianz mittels Genotyping-by-Sequencing (GBS) abgezielt, um somit die Selektion auf komplexe Merkmale zu ermöglichen (Bhat et al., 2016). Jedoch ist dies durch das Fehlen von vollständigen Referenzgenomen und durch den hohen Grad an Heterozygotie bei Hopfen zurzeit nicht möglich. Die hier vorgestellte Aufklärung der Genstruktur der PT1 und der PT2 ist die Vorarbeit für die gezielte

Sequenzierung der genregulatorischen Bereiche der PT1 und PT2 ohne vollständiges Referenzgenom.

## 2. Material und Methoden

Die Basis für die Genstrukturaufklärung bilden funktional geprüfte Sequenzen der Prenyltransferase 1 (PT1, GenBank-Akzession KM222441) und Prenyltransferase 2 (PT2, GenBank-Akzession KM222442). Diese Sequenzen wurden genutzt, um mittels BLAST-Algorithmus ([blast.ncbi.nlm.nih.gov](http://blast.ncbi.nlm.nih.gov)) die genomischen Regionen zu identifizieren, auf welchen die PT-Gene lokalisiert sind. Hierzu wurden die Genome der Hopfensorten (*Humulus lupulus* L.) „Tea Maker“ (TMK), „Shinzu Wase“ (SSW) und später „Cascade“ (CAS) genutzt (HopBase, [www.hopbase.org](http://www.hopbase.org), Stand 15.02.2019). Die so erhaltenen Scaffolds (SSW, TM) und Contigs (CAS) wurden dann anhand der Zahl der vorhandenen PT-Exons und anhand der Qualität des BLAST-Alignments geprüft. Die teilweise lückenhaften Scaffolds von SSW und TMK wurden im Bereich vor dem jeweiligen Exon 1, der wahrscheinlichen Promotorregion, vervollständigt. Die TMK- und SSW-Scaffolds auf denen die PT1-Exons gefunden wurden, ließen sich bis zur Veröffentlichung des CAS-Genoms (Februar 2019) nicht ohne weitere Sequenzierungen vervollständigen. Daher wurde ergänzend zur Sequenzuntersuchung eine Neusequenzierung, basierend auf PCR und Sanger-Sequenzierung, zur Aufklärung der Genstruktur durchgeführt.

Die genomische DNA für die PCR-Analysen wurde aus Pflanzenmaterial der Sorte „Polaris“ des Hopfenforschungszentrums Hüll (Wolnzach, Deutschland) gewonnen. Die DNA-Extraktion wurde basierend auf dem Protokoll von Fleischer et al. (2004) durchgeführt. Für die PCR-Analysen wurde eine Vielzahl von Primern entwickelt, um die Sequenzlücken vorhandener Scaffolds zu ergänzen. Hierfür wurden PCRs und nested PCRs durchgeführt. Falls möglich, wurde direkt sequenziert (Eurofins Scientific, Luxemburg). In einigen Fällen war vor der Sequenzierung die Aufreinigung (Biolone Isolate II PCR & Gel Kit, UK) oder die Klonierung des PCR-Fragments notwendig (Zero Blunt™ PCR Cloning Kit, Invitrogen, USA). Die so erhaltenen Sequenzen wurden mit dem De-Novo-Assembly Tool aus Geneious R11 (Geneious 11.1.5, Neuseeland) mit den ursprünglichen Scaffolds assembliert. In Fällen, wo die gesamten Transkriptsequenzen der PT-Gene auf einem Scaffold oder Contig vorhanden waren, konnten Alignments mit dem Programm SPLIGN unter Berücksichtigung wahrscheinlicher Splice Stellen durchgeführt werden (Kapustin et al., 2008). In Fällen, wo aufgrund von Sequenzlücken SPLIGN nicht angewendet werden konnte, wurde ein MAFFT-Alignment mit anschließender Splice-Stellen-Analyse mit dem NetPlantGene Programm durchgeführt (Hebsgaard et al., 1996).

Die so erhaltene erste Version der Genstruktur wurde mittels MAFFT-Alignments mit veröffentlichten Transkriptsequenzen abgeglichen (Clark et al., 2013; Natsume et al., 2015). Unabhängig davon wurden zur Identifizierung von Sequenzmustern und Transposons die Sequenzen mit dem Programm CENSOR (<https://www.girinst.org/censor/index.php>; Datenbank RepBase Version 23.12) analysiert (Kohany et al., 2006). Im letzten Schritt wurden die Promotorbereiche mit dem Genomatix ModellInspector (Genomatix, Deutschland; Promotor Module Library Version 6.2) geprüft (Klingenhoff et al., 1999).

## 3. Ergebnisse

Insgesamt wurden bei der Datenbankabfrage der HopBase 11 Sequenzen identifiziert, die mit sehr hoher Sequenzübereinstimmung einzelne Exons der Prenyltransferasen enthalten (Daten nicht gezeigt). Das PT1 Gen konnte auf keinem Scaffold vollständig identifiziert werden. Daher wurden die Scaffolds ausgewählt, die möglichst wenig Sequenzlücken hatten und gleichzeitig viele PT1-Exons enthielten (SSW scaffold\_75308, TMK scaffold12808). Die Sequenzlücke vor dem Exon 1 wurde mittels PCR-Sangersequenzierung geschlossen. Durch nested PCRs konnte gezeigt werden, dass zwischen Exon 3 auf scaffold12808 und Exon 4 auf scaffold\_75308 ein 7 kb langes Intron liegt und diese Sequenzen somit zum gleichen Gen gehören (Abbildung 1 A). Die Exons der PT2 konnten auf dem TMK Scaffold scaffold15836 gefunden werden (Abbildung 1 B). Anhand von Alignments mit dem im Februar 2019 veröffentlichten CAS-Genom zeigte sich, dass beide PTs auf dem CAS Contig 006344F in einer Entfernung von 170 kb zueinander lokalisiert sind (Daten nicht gezeigt). Aufgrund guter Qualität (keine Sequenzlücken, lange Sequenzen) basieren die weiteren bioinformatischen Analysen auf dem Contig 006344.



Abbildung 1. MAFFT-Alignment des Contigs 006344F mit den Scaffolds scaffold12808, scaffold\_75308 und scaffold15836 (graue Linien mit schwarz markierten Nichtübereinstimmungen). Die Exons der Gene (A) PT 1 und (B) PT 2 sind durch schwarze Pfeile repräsentiert. PCR-Fragmente und mit Sanger komplementierte Sequenzabschnitte sind grün hervorgehoben. Das Suffix \_Rev zeigt Sequenzinvertierung an.

Die Ergebnisse der SPLIGN-Analyse zeigten, dass beide Prenyltransferasen aus 10 Exons bestehen, die einen Bereich von 7 kb beim PT1-Gen und von 31 kb beim PT2-Gen umfassen (Abbildung 2). Dabei ist die Länge des PT2-Gen vor allem auf das 18 kbp lange PT2-Intron 4 zurückzuführen. Die MAFFT-Alignments zwischen den Genom- und den Transkriptsequenzen bestätigten die Exonstruktur und zeigten zudem, dass es vor und nach beiden PT-Genen nicht-translatierte Regionen (UTR) gibt (Abbildung 2). Die Untersuchung der Sequenz mit CENSOR ergab, dass vor allem im Intron 4 der PT2 zwei längere (>2 kb) Long Terminal Retrotransposons (LTR) vorliegen. Der Vergleich der PT-Promotorenregionen zwischen CAS, SSW und TMK zeigte, dass die regulatorischen Regionen der PT-Gene zu 99% übereinstimmen und somit nur wenig genetische Varianz aufweisen. Diese Regionen wurden durch das Programm Genomatix ModellInspector

anhand einer Datenbank auf experimentell bestätigte und zusammengehörige Gruppen von Transkriptionsfaktorbindungsstellen, den sog. TFBS-Modulen, geprüft (Klingenhoff et al., 1999). Im Fall der PT1 wurde im Bereich 1kb vor Exon 1 nur ein Modul mit DNA-binding-with-one-finger und homeo-domain-leucine-zipper Motiven gefunden (Schneidereit et al., 2008), welches bei Arabidopsis die gewebespezifische Genexpression steuert. Im Fall der PT2 wurden im selben Bereich zwei weitere aus Arabidopsis bekannte Module gefunden, ein trockenstress-aktivierbares Modul bestehend aus einer GT-Box und einer Myc-Erkennungssequenz (Simpson et al., 2003) und ein brassinosteroid-assoziiertes Modul mit einer Myb-Erkennungssequenz und einer E-box (Li et al., 2009).

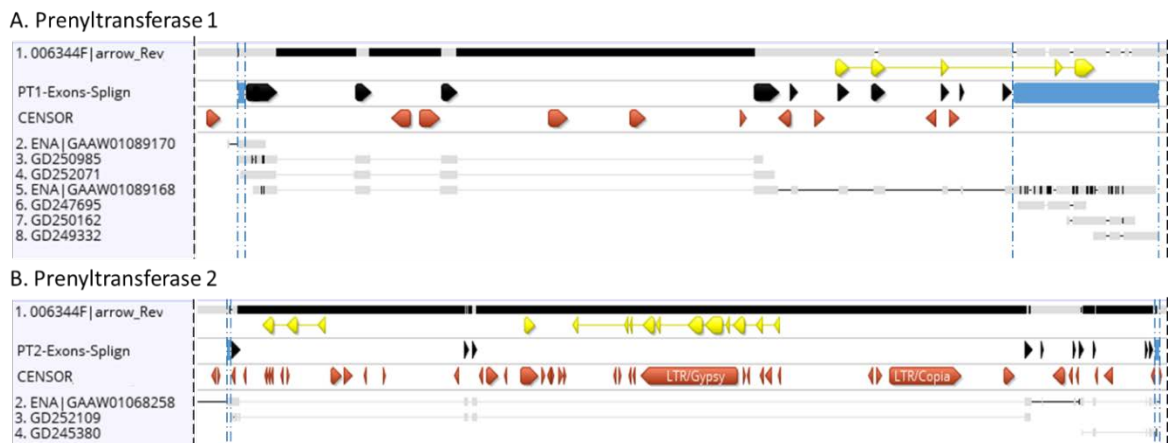


Abbildung 2. MAFFT-Alignment des Contigs 006344F mit einer Auswahl an Transkriptsequenzen (graue Linien mit schwarz markierten Nichtübereinstimmungen). Exons der Gene (A) Prenyltransferase 1 und (B) Prenyltransferase 2 sind durch schwarze Pfeile repräsentiert. Bereiche vor und nach den Exons wurden als UTR blau hervorgehoben. Das Suffix \_Rev zeigt Sequenzinvertierung an.

#### 4. Diskussion

Die 63%-ige Übereinstimmung der PT-Transkripte sowie die Gemeinsamkeiten der Genstruktur (je 10 Exons flankiert von ähnlich langen 5'- und 3'-UTRs) deuten darauf hin, dass die beiden PT-Gene miteinander verwandt sind. Zudem haben beide PT-Gene vergleichsweise lange Introns von mehr als 2 kb, welches bei weniger als 2% aller pflanzlichen Gene vorkommt (Li et al., 2017). Dabei ist das PT2-Intron 4 mit 18 kb ungewöhnlich lang, da beispielsweise bei Arabidopsis bislang nur 16 Introns identifiziert wurden, die länger als 5 kb sind (Chang et al., 2017). Die Ursachen für die Länge der PT-Introns lassen sich vermutlich auf die Aktivität von Retrotransposons, insbesondere von LTRs zurückführen, die bei Hopfen 32% des Gesamtgenoms (Natsume et al., 2015) und im Bereich der PT-Gene 21% ausmachen. Allein die zwei längsten LTRs, Gypsy und Copia, im PT2-Intron 4 machen 31% der Intronlänge aus. Inwieweit die langen Introns die Genfunktionalität beeinflussen bedarf weiterer Untersuchungen. Vorläufige Analysen zeigen jedoch, dass auch in den Introns regulatorische Elemente enthalten sind (Daten nicht gezeigt). In jedem Fall sind wichtige regulatorische Bereiche in der Promotorregion zu erwarten, weshalb hier gezielt nach experimentell belegten TFBS-Modulen gesucht wurde (Klingenhoff et al., 1999). Für die PT1 wurden ein und für die PT2 zwei plausible Module

identifiziert. Es ist davon auszugehen, dass weit mehr TFBS-Module vorhanden sind, doch enthält die zugrundeliegende Datenbank zurzeit erst 111 pflanzliche TFBS-Module. Die gefundenen Module waren im Bereich der ersten 400 bp vor dem jeweiligen Exon 1 lokalisiert. Aus diesem Grund werden sich die geplanten gezielten Sequenzierungen auf diese Regionen konzentrieren.

## 5. Schlussfolgerung

Anhand von drei unterschiedlichen Hopfengenomen und mithilfe weiterer Transkriptdaten konnte eine erste vorläufige Genstruktur der PT1 und PT2 ausgearbeitet werden. Aufgrund der hohen Ähnlichkeit vor allem im Promotorbereich und sich wiederholender Genstruktur zwischen den beiden Genen kann davon ausgegangen werden, dass es sich hier nicht um Pseudogene, sondern um funktionale Gene handelt. Für die endgültige Abklärung müssten jedoch (1) ein vollständiges Hopfengenom vorliegen und (2) funktionelle Experimente mit den Genen und Genprodukten durchgeführt werden. Trotzdem sind diese Genstrukturen eine plausible Grundlage für die Identifizierung potentiell regulatorischer Elemente für eine gezielte Sequenzierung. Sollten die zukünftig in diesen Bereichen gefundenen Sequenzvariationen mit dem Gehalt der Bittersäuren korrelieren, wäre ein weiteres starkes Indiz für die Richtigkeit der vorgestellten Genstruktur gefunden.

## 6. Literatur

Bhat, J.A., Ali, S., Salgotra, R., Mir, Z., Dutta, S., Jadon, V., Tyagi, A., Mushtaq, M., Jain, N., Singh, P.K., ... Prabhu, K.V. (2016). Genomic selection in the era of next generation sequencing for complex traits in plant breeding, *Frontiers in Genetics*, 7 Artikel 221.

Chang, N., Sun, Q., Hu, J., An, C., Hangbo, G. (2017). Large introns of 5 to 10 kilo base pairs can be spliced out in arabidopsis. *Genes*, 8(8) Artikel 200.

Clark, S.M., Vaitheeswaran, V., Ambrose, S., Purves, R.W., Page, J. (2013). Transcriptome analysis of bitter acid biosynthesis and precursor pathways in hop. *BMC Plant Biology*, 13(1) Artikel 12.

Fleischer, R., Horlemann, C., Schwekendiek, A., Kling, C., Weber, G. (2004). AFLP fingerprinting in hop. *Genetic Resources and Crop Evolution*, 51(2): 211–220.

Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouzé, P., Brunak, S. (1996). Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. *Nucleic acids research*, 24(17): 3439–52.

Kapustin, Y., Souvorov, A., Tatusova, T., Lipman, D. (2008). Splign: algorithms for computing spliced alignments with identification of paralogs. *Biology Direct*, 1 Artikel 20.

Klingenhoff, A., Frech, K., Quandt, K., Werner, T. (1999). Functional promoter modules can be detected by formal models independent of overall nucleotide sequence similarity. *Bioinformatics*, 15(3): 180–186.

Kohany, O., Gentles, A.J., Hankus, L., Jurka, J. (2006). Annotation, submission and screening of repetitive elements in Repbase. *BMC bioinformatics*, 7 Artikel 474.

Li, Y., Xu, Y., Ma, Z. (2017). Comparative analysis of the exon-intron structure in eukaryotic genomes, *Yangtze Medicine*, 01: 50–64.

Natsume, S., Takagi, H., Shiraishi, A., Murata, J., Toyonaga, H., Patzak, J., Takagi, M., Yaegashi, H., Uemura, A., Mitsuoka, C., ... Eiichiro O (2015). The Draft Genome of Hop (*Humulus lupulus*), an Essence for Brewing. *Plant & cell physiology*, 56(3): 428–41..