



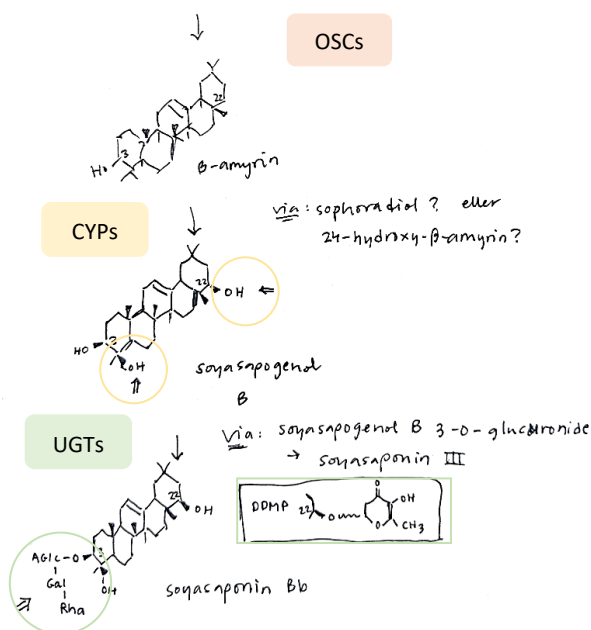
Saponiner i ärt

– Biosyntes av soyasaponiner och dess genetiska mekanismer i ärtväxter

Saponins in peas

– *Biosynthesis and genetic regulation of soyasaponins in legumes*

Ines Falk Tingström



Självständigt arbete i biologi - 15 hp

Sveriges lantbruksuniversitet, SLU

Institutionen för växtförädling

Trädgårdsingenjör: odling – kandidatprogram

Alnarp 2022

Saponiner i ärt – Biosyntes av soyasaponiner och dess genetiska mekanismer i ärtväxter

Saponins in peas – Biosynthesis and genetic regulation of soyasaponins in legumes

Ines Falk Tingström

Handledare: Cecilia Hammenhag, SLU, Institutionen för växtförädling

Examinator: Ida Lager, SLU, Institutionen för växtförädling

Omfattning: 15 hp

Nivå och fördjupning: Grundnivå, G2E

Kurstitel: Självständigt arbete i biologi

Kurskod: EX0855

Program/utbildning: Trädgårdsingenjör: odling – kandidatprogram

Kursansvarig inst.: Institutionen för biosystem och teknologi

Utgivningsort: Alnarp

Utgivningsår: 2022

Omslagsbild: Skiss över saponinsyntes, författarens egen

Nyckelord: *Pisum sativum*, baljväxter, Fabaceae, proteingrödor, växtförädling, jämförande genomik.

Sveriges lantbruksuniversitet

Fakulteten för landskapsarkitektur, trädgårds- och växtproduktionsvetenskap

Institutionen för växtförädling

Publicering och arkivering

Godkända självständiga arbeten (examensarbeten) vid SLU publiceras elektroniskt. Som student äger du upphovsrätten till ditt arbete och behöver godkänna publiceringen. Om du kryssar i **JA**, så kommer fulltexten (pdf-filen) och metadata bli synliga och sökbara på internet. Om du kryssar i **NEJ**, kommer endast metadata och sammanfattning bli synliga och sökbara. Fulltexten kommer dock i samband med att dokumentet laddas upp arkiveras digitalt.

Om ni är fler än en person som skrivit arbetet så gäller krysset för alla författare, ni behöver alltså vara överens. Läs om SLU:s publiceringsavtal här: <https://www.slu.se/site/bibliotek/publicera-och-analysera/registrera-och-publicera/avtal-for-publicering/>.

☒ JA, jag/vi ger härmed min/vår tillåtelse till att föreliggande arbete publiceras enligt SLU:s avtal om överlåtelse av rätt att publicera verk.

☐ NEJ, jag/vi ger inte min/vår tillåtelse att publicera fulltexten av föreliggande arbete. Arbetet laddas dock upp för arkivering och metadata och sammanfattning blir synliga och sökbara.

Sammanfattning

Efterfrågan på proteinrika vegetabiliska livsmedel ökar och intresse finns från både konsument, politiker och industri att använda svenskodlade råvaror i större utsträckning. Ärtor (*Pisum sativum* L.) har en lång odlingstradition i Sverige och ett högt proteininnehåll, men innehållet av soyasaponiner begränsar användningen då de ger en bitter smak till livsmedel.

Denna litteraturstudie är en del i forsknings- och växtförädlingsprojektet "Ärtan – Garantin för framtidens gröna protein" inom SLU Grogrund, med syftet att på sikt ta fram nya ärtsorter för att möta efterfrågan på marknaden. Genom att studera befintlig litteratur besvaras frågan om vad som är känt om mekanismer bakom soyasaponininnehållet och -variationen i ärtor. Studien redogör även för vad som finns dokumenterat angående soyasaponinsyntes i besläktade och närmre studerade arter. Genom att tillämpa jämförande genomik på tillgängliga sekvensdata för ärtor och närbesläktade arter undersöks förekomsten av homologa sekvenser.

Soyasaponiner är saponiner av oleanan-typ och förekommer i flera arter inom familjen ärtväxter. Ett tidigt steg i soyasaponinsyntesen är cykliseringen av 2,3-oxidosqualene till β -amyrin, för vilket ett multifunktionellt- och ett specifikt enzym tidigare identifierats i ärtor. Även genen för det specifika β -amyrin-syntaset har lokaliserats och uttrycksmönstret för denna samt för ackumuleringen av soyasaponiner har redogjorts för. Övriga steg i syntesen är hydroxylering av β -amyrin till aglykonen soyasapogenol B följt av inbindning av socker och andra molekyler till aglykonen, vilka i andra växter påvisats involvera enzym inom grupperna cytochrom P450-beroende monooxygenaser respektive glykosyltransferaser. Information om dessa grupper av enzym och andra relaterade gener, samt vad som är känt angående dess funktion sammanfattas för att underlätta förståelsen för hur motsvarande, hittills utforskade, mekanismer i ärtor skulle kunna se ut.

Jämförande genomik är ett verktyg för identifiering av homologa gener och i studien identifieras flera protein och sekvenser i ärtgenomet som potentiellt kan kopplas till syntesen av soyasaponiner. Metoder för vidare homologianalys och validering av funktionen hos potentiella gener i ärtor diskuteras. Variation i soyasaponininnehåll har konstaterats som primärt genotypberoende vilket gör det relevant, men är även en förutsättning för, att vidare undersöka ärtas genom och genetik i syfte att förbättra proteinråvarans smakkvalitet.

I syfte att uppnå de nationella- och globala hållbarhetsmålen med en effektiviserad livsmedelsproduktion och en ökad andel vegetabiliska livsmedel kan baljväxter vara en del av lösningen, och växtförädling spela en avgörande roll.

Nyckelord: *Pisum sativum*, baljväxter, Fabaceae, proteingrödor, växtförädling, jämförande genomik.

Abstract

The demand for vegetarian and protein rich foods are increasing, and to use locally produced crops are in the interest of both consumers, politicians, and the processing industry. In Sweden, peas (*Pisum sativum* L.) have been a widely grown crop and staple food throughout the history of farming and are high in protein. Due to the bitter taste of soyasaponins present in peas, the use of food products based on peas and pea protein is limited.

This literature study is a part of a research and plant breeding project on peas within SLU Grogrund, a competence center and knowledge hub for Swedish plant breeding, with a long-term goal to make new pea cultivars with desired traits available on the market. The question on what is known about genetic regulation of soyasaponins in peas and other factors affecting the saponin content is answered by studying existing literature. The report also points out what is published on more well studied soyasaponin producing species and their known genes and proteins involved in the soyasaponin biosynthesis. By comparative genomics the occurrence of homolog sequences in peas is analyzed.

Soyasaponins are oleanane-type saponins found in many leguminous plant species. In the biosynthesis, an early step is the cyclization of 2,3-oxidoqualene to β -amyrin for which one multi-functional and one specific enzyme previously has been identified in peas. The gene encoding the specific β -amyrin-synthase has been localized and the pattern for expression of this gene and for the accumulation of soyasaponins have been reported on. Following steps of the synthetic pathway is hydroxylation of β -amyrin to the soyasapogenol B aglycone and the binding of sugars and other moieties to the aglycone structure, for which no enzymes in peas have been identified. In other leguminous species these reactions have been shown to involve enzymes categorized as cytochrome P450-dependent monooxygenases and glycosyltransferases, respectively. Available information on these groups of enzymes, other genes of interest, and their known characteristics are summarized to spread some light on how equivalent, and to date undocumented, mechanisms in peas could look like.

Comparative genomics is a tool for identification of gene homologs, and in this study several proteins and sequences in the pea genome are identified as potentially involved in the soyasaponin biosynthetic pathway. Methods for further analysis of homologs, and for validating the function of potential genes in peas are being discussed. Variation has been proven to be mainly genotype-dependent, making it of interest to, but is also a prerequisite for, searching the pea genome and genetics in the aim of improving the taste of the pea protein product.

To achieve the sustainability goals set by national and global stakeholders and to make the food production more resource efficient, plant breeding on legumes could play a main role.

Keywords: *Pisum sativum*, pulses, Fabaceae, protein crops, plant breeding, comparative genomics.

Förord

Jag vill säga tack. Tack till min handledare Cecilia, för att du uppmuntrade mig att ta mig an detta arbete, varit ett bollplank genom processen och inte tvekat att komma med feedback och respons. Tack till mina sambos i Arlöv och Göteborg, för att ni peppat och inspirerat, hållit mig sällskap, bryggt kaffe och lagat mat. Jag hade möjligtvis tagit mig igenom detta även utan er, men vägen hit hade varit bra mycket svårare.

Ines

Innehållsförteckning

Tabellförteckning	9
Figurförteckning	10
Förkortningar	11
Förtydliganden	12
1. Inledning	14
1.1. Bakgrund	14
1.1.1. En värld i förändring – nationella och globala mål	14
1.1.2. Ärtas genetik	15
1.1.3. Rollen som proteingröda	16
1.2. Syfte	17
1.3. Frågeställning	17
2. Metod	18
2.1. Litteratursökning	18
2.2. Jämförande genomik	18
2.3. Avgränsningar	19
3. Resultat	20
3.1. Saponiner – klassificering, kemi och biosyntes	20
3.2. Enzym kopplade till soyasaponinsyntes i ärtväxter	23
3.2.1. Oxidosqualene-cyklaser	24
3.2.2. Cytokrom P450-beroende monooxygenaser	27
3.2.3. Glykosyltransferaser	29
3.3. Variation i saponinhalt	31
3.3.1. Faktorer som styr variationen	32
3.3.2. Transkriptionsfaktorer	33
4. Diskussion	34
4.1. Genetiska mekanismer bakom saponininnehåll i ärt	34
4.2. Litteratursökning och jämförande genomik som metod	35
4.3. Tolkning av resultat från jämförande genomik	36
4.4. Växtförädling av sorter med lägre saponininnehåll	40
4.5. Slutsats	43
Referenser	44
Bilagor	58

Tabellförteckning

Tabell 1: Urval av träffar från sökning i BLAST med cytokrom P450-beroende C-24-hydroxylas ur underfamiljen CYP93E.	28
Tabell 2: Urval av träffar från sökning i BLAST med cytokrom P450-beroende hydroxylas ur underfamiljen CYP72A.....	29
Tabell 3: Urval av träffar från sökning i BLAST med glykosyltransferaser involverade i syntesen av soyasaponin Bb och DDMP-saponin i soja.	31
Tabell 4: Urval av träffar från sökning i BLAST med transkriptionsfaktorer vilka uppreglerar uttryck av gener involverade i soyasaponinsyntes.....	33
Tabell B1: Sammanställning av enzym som kan kopplas till syntesen av soyasaponiner eller ämnen soyasaponiner kan erhållas ur.....	58
Tabell B2: Sammanställning av resultat från sökning i BLAST, i databaser tillgängliga via NCBI Nucleotide Collection respektive Whole Genome Shotguns.....	60

Figurförteckning

Figur 1: Produktion i 10 000 ton för några vanligen odlade baljväxtgrödor, i Sverige, i Europa och globalt. Baserat på data för år 2017 från FAO (2021).	14
Figur 2: Fylogenetiskt träd med släkten inom Fabaceae samt <i>Arabidopsis</i> som referens. Skapad i verktyget PhyloT via Tree of Life baserat på data från NCBI Taxonomy Database.	16
Figur 3: Visualiserad och förenklad syntesväg för 2,3-oxidosqualene från acetyl-koenzym A.	20
Figur 4: I soja föreslagen syntesväg och skelettformler för ämnena 2,3-oxidosqualene, β -amyrin, soyasapogenol B och DDMP-saponin.	21
Figur 5: Skelettformel för soyasaponin A samt soyasaponin B och DDMP-saponin, och omvandlingen däremellan.	22
Figur 6: Saponininnehåll i ärtfröer av sorten 'Caméor' under fröutveckling. Figur erhållen från Vernoud <i>et al.</i> (2021) och översatt från engelska med tillstånd från Oxford University Press.....	23
Figur 7: Cyklisering av 2,3-oxidosqualene till β -amyrin. Skelettformler för substrat och produkt samt oxidosqualene-cyklasers karaktäriseristiska aminosyrasekvenser.	24
Figur 8: Relativ nivå av uttryckt mRNA för flera förmodade oxidosqualene-cyklaser under fröutveckling i ärtsorten 'Caméor'. Figur erhållen från Vernoud <i>et al.</i> (2021) och översatt från engelska med tillstånd från Oxford University Press.	26
Figur 9: Föreslagen syntesväg av hydroxylering från β -amyrin till soyasapogenol B... ..	27
Figur 10: Föreslagen syntesväg i soja (<i>Glycine max</i>) från soyasapogenol B till DDMP-saponin.	30

Förkortningar

Acc.nr.	accessionsnummer
bHLH	basic Helix-Loop-Helix transkriptionsfaktor
BLAST	Basic Local Alignment Search Tool
Blastn	Standard Nucleotide BLAST
BR	brassinosteroider
C	kol (i molekylstrukturer, numrerade)
cDNA	komplementärt DNA (<i>eng.</i> complementary DNA)
CSyGTs	cellulosasyntas-erhållet glykosyltransferas
CYPs	cytokrom P450-beroende monooxygenaser
Dap	dagar efter pollinering (<i>eng.</i> days after pollination)
DDMP	2,3-dihydro-2,5-dihydroxy-6-methyl-4H-pyran-4-one
DNA	deoxyribonukleinsyra (<i>eng.</i> deoxyribonucleic acid)
Gbp	Giga-baspar
GTs	glykosyltransferaser
mRNA	budbärar-RNA (<i>eng.</i> messenger RNA)
NC	NCBI Nucleotide Collection
NCBI	National Center for Biotechnology Information
OSCs	oxidosqualene-cyklaser
PCR	Polymerase Chain Reaction
PSPG-box	Plant Secondary Product Glycosyltransferase box
RNA	ribonukleinsyra (<i>eng.</i> ribonucleic acid)
Tblastn	Translated BLAST
TFs	transkriptionsfaktorer
UDP	uridindifosfat
UGTs	uridindifosfat-beroende glykosyltransferaser
WGS	Whole-genome Shotguns

Begrepp och förtydliganden

Accession	tillskott till en kollektion, exempelvis sekvensdata till en databas eller frömaterial till ett förädlingsprojekt.
Accessionsnummer	ett nummer som identifierar accessioner i sekvensdatabaser.
DDMP-saponin	en grupp soyasaponiner. I arbetet avses det i ärtor förekommande DDMP-saponinet Soyasaponin βg. Synonymer i litteraturen är Soyasaponin VI och Chromosaponin I.
Homologa	sekvenser av gemensamt evolutionärt ursprung, vilka divergerat på grund av artbildning (ortologa) eller genduplikation (paraloga). Kan ha motsvarande funktion bevarad.
Homozygot	en individ med enbart likadana (homo-) alleler av en gen, till skillnad från heterozygot vilket syftar till en individ med olika (hetero-) alleler av en gen.
Hydrolysering	kemisk process där en molekyl klyvs efter addering av en vattenmolekyl.
Hydroxylering	kemisk reaktion, inbindning av en hydroxylgrupp (-OH) till en kemisk förening.
Inducerad mutagenes	avsiktligt framkallande av mutationer
Kopplingsgrupp	samling gener som är tillräckligt nära varandra för att med stor sannolikhet nedärvas tillsammans. I helgenomsekvenseringar motsvarar de ofta kromosomerna.
Monofyletisk grupp	alla nu levande organismer som kan härledas till en gemensam anfader.

N-terminal / C-terminal	N-terminal är den ände av ett protein som bildas först under translationen, bestående av en fri aminogrupp ($-\text{NH}_2$). C-terminal är motsatt ände, bestående av en fri karboxylgrupp ($-\text{COOH}$). Sekvenser skrivs vanligen från N- till C-terminalen.
Sekundärmetabolit	metabolisk produkt som inte krävs för de direkt livsuppehållande systemen, som tillväxt, utveckling och reproduktion.
Soyasaponin Bb	grupp B-saponin påträffad i ärtor. Synonymer i litteraturen är Soyasaponin I, ibland även kort saponin B.
Soyasaponiner	triterpenoida glykosida saponiner som är naturligt förekommande i soja och andra ärtväxter.
Transgenteknik	tekniker som innebär att arvs massa överförs horisontellt mellan organismer, i vardagligt tal genmodifiering. Kan vara inom eller mellan arter.
Transposon	genssekvens vars funktion är att byta plats i genomet där den förekommer. Även kallad mobilt genetiskt element eller hoppande gen.

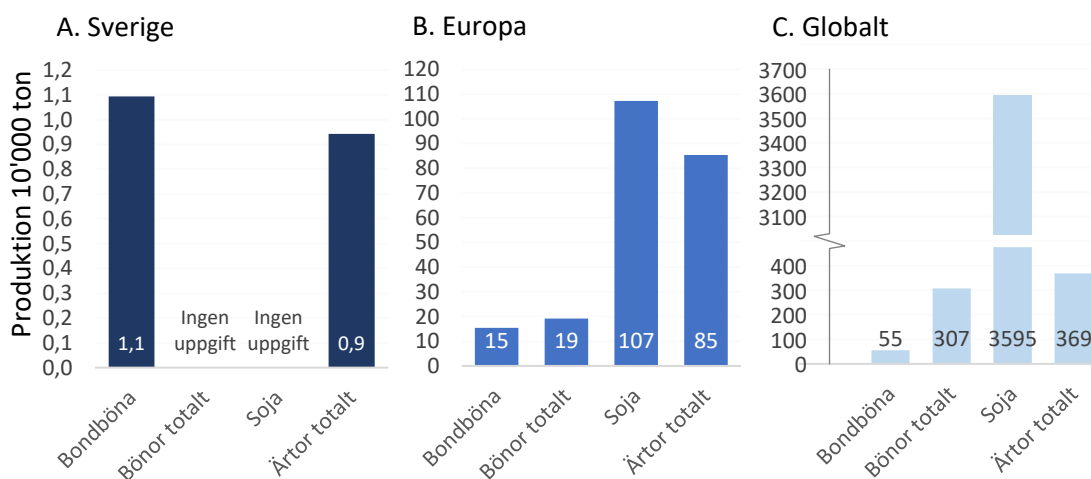
1. Inledning

1.1. Bakgrund

1.1.1. En värld i förändring – nationella och globala mål

År 2015 antog Sverige som en av 193 medlemsstater i Förenta nationerna resolutionen Agenda 2030 och de 17 globala målen för hållbar utveckling (United Nations 2015). En övergång till en större andel vegetabiliska proteinkällor efterfrågas från politiskt håll i syfte att leva upp till ställda hållbarhetsmål och är en del av den livsmedelsstrategi som regeringen gjort gällande till år 2030 (Regeringen 2017). Enligt en översiktsstudie av Boye *et al.* (2010) efterfrågar även livsmedelsproducenter och konsumenter livsmedel med lägre koldioxidavtryck, och vidare föreslås baljväxter som ett lämpligt alternativ till animalier.

Baljväxter är arter inom familjen ärtväxter (Fabaceae), där exempelvis soja (*Glycine max* L.), bönor (*Phaseolus* spp.), lupin (*Lupinus* spp.), bondböna (*Vicia faba* L.), lins (*Lens culinaris* Medik.) och ärt (*Pisum sativum* L.) är vanligt odlade proteingrödor (Figur 1) (De Ron *et al.* 2017; Food and Agriculture Organization of the United Nations (FAO) 2021).



Figur 1: Produktion i 10 000 ton för några vanligen odlade baljväxtgrödor (A) i Sverige, (B) Europa och (C) globalt. Bondböna (*Vicia faba*), gröna och torkade bönor (*Phaseolus* spp.), soja (*Glycine max*) samt gröna och torkade ärtor (*Pisum sativum*). Baserat på data för år 2017 från FAO (2021).

Baljväxternas alla delar är ätliga men i litteraturen åsyftas främst fröerna, i synnerhet i fråga om humankonsumtion. En stor andel av de proteingrödor som produceras idag går till animalieproduktion (Jordbruksverket 2021; Semba *et al.* 2021), och en betydande del

av det ursprungliga proteinet går därmed förlorat på vägen till slutkonsument (Shepon *et al.* 2016).

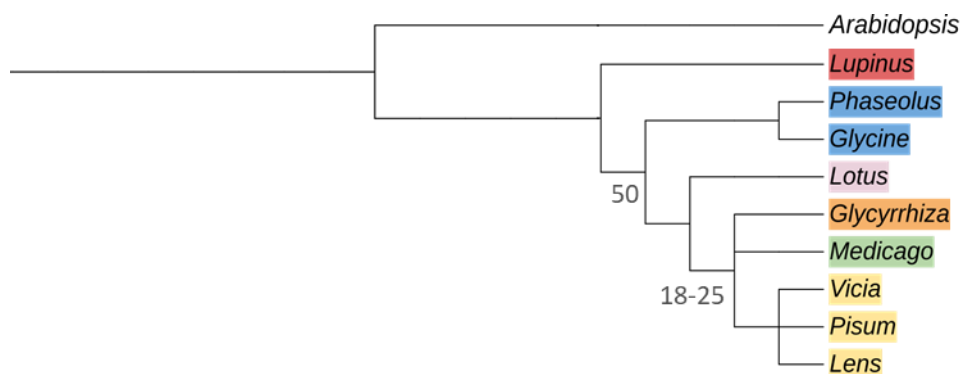
Många arter inom Fabaceae kan ingå symbios med kvävefixerande jordbakterier, vanligen inom släktet *Rhizobium*, genom att inducera bildandet av rotnölar, noder, där bakterierna förser värdväxten med växttillgängligt kväve (Stacey 2007). Att öka andelen baljväxter inom den europeiska jordbruksproduktionen anses ha flera potentiella fördelar, bland annat odlingstekniskt genom att diversifiera en i övrigt spannmålsdominerad växtföljd, minska kvävebehovet och -förlusterna samt bidra med andra ekosystemtjänster, men även för att minska beroendet av importerad soja och sojabaserade produkter (Reckling *et al.* 2016; Zander *et al.* 2016; Watson *et al.* 2017).

Enligt Zander *et al.* (2016) är den generella trenden att andelen baljväxter inom det europeiska jordbruket minskat de senaste decennierna och att växtförädling och nya växtförädlingsprogram krävs för att vända utvecklingen. SLU Grogrund är ett centrum som samordnar växtförädlingskompetens inom akademi och industri, vilket inrättats enligt beslut från Sveriges regering i syfte att uppnå målen inom livsmedelsstrategin kring en säkrad tillgång på växtsorter (Regeringen 2019; SLU Grogrund 2021a). Detta för att möjliggöra hållbarhet och konkurrenskraft inom den nationella jordbruks- och trädgårdsproduktionen (Regeringen 2019).

1.1.2. Ärtas genetik

Ärta domesticerades i mellanöstern i samband med den neolitiska revolutionen för cirka tiotusen år sedan (Kreplak *et al.* 2019), och har en lång odlingstradition även i Sverige som en av de första odlade grödorna (Leino & Nygård 2008). Underarten *Pisum sativum* subsp. *sativum* innefattar bland andra gröna- och gula spritärter vilka är viktiga grödor inom livsmedelsindustrin (Nationalencyklopedin u.å.-a). Arten är diploid med sju kromosompar ($2n = 14$) och genomets storlek uppskattas till cirka 4,45 Gbp (Doležel *et al.* 1998; Kreplak *et al.* 2019). Enbart 104 av knappt 700 arter inom Fabaceae har större genom än ärta, varav alla förutom tre tillhör stammen Fabaceae (Kreplak *et al.* 2019).

Redan Mendel (1866) använde ärta som modellväxt för sina försök som resulterade i formulerandet av genetikens lagar, men först år 2019 publicerades ett referensgenom för ärta vilket beräknas representera cirka 88% av genomets fulla sekvens (Kreplak *et al.* 2019). På grund av genomets storlek har enligt Kreplak *et al.* (2019) artspecifika verktyg för DNA-baserad växtförädling, som genetiska markörer för selektion och alleler kopplade till fenotyp, dröjt. Detta arbete har kommit längre för arter som exempelvis soja och modellväxten tornlusern (*Medicago truncatula* Gaertn.), vilka båda har mindre genom (Schmutz *et al.* 2010; Young *et al.* 2011; Kreplak *et al.* 2019). Projektet att sekvensera det 13 Gbp stora genomet från för ärta närbesläktade bondböna är påbörjat men ännu inte avslutat (Carrillo-Perdomo *et al.* 2020).



Figur 2: Fylogenetiskt träd med släkten inom Fabaceae samt *Arabidopsis* (Brassicaceae, korsblommiga växter) som referens. Siffror representerar uppskattad tid sedan divergens i miljoner år. Släkten med samma färgmarkering tillhör samma stam. Röd: *Lupinus* (Genisteae); blå: *Phaseolus* och *Glycine* (Phaseoleae); rosa: *Lotus* (Loteae); orange: *Glycyrrhiza* (Galegeae); grön: *Medicago* (Trifolieae); gul: *Vicia*, *Pisum* och *Lens* (Fabeae). Skapad i verktyget PhyloT via Tree of Life (<https://phylot.biobyte.de/index.cgi> hämtad 2021-12-20) baserat på data från NCBI Taxonomy Database.

Ärta, bondböna och tornlusern tillhör samma morfologiska grupp, vilken divergerade från gruppen som bland annat innefattar soja och bönsläktet (*Phaseolus*) för cirka 50 miljoner år sedan (Figur 2) (Kreplak *et al.* 2019). Ärta har flest släktspecifika gener, cirka 33% av samtliga gener, och högst helgenommutationstakt bland alla studerade arter inom Fabaceae och har en hög proportion nyligen uppkomna paraloger i jämförelse med tornlusern, vilket associeras med genomets expansion och stammens divergens från Trifolieae för cirka 25 till 18 miljoner år sedan (Kreplak *et al.* 2019). Ungefär 83% av ärtans genom består av repetitiva element, varav de flesta utgörs av transposoner och i synnerhet retrotransposoner (Kreplak *et al.* 2019). Sekvenserna hos repetitiva element i tornlusern och ärta är likartade, om än betydligt längre i genomet hos den senare (Macas *et al.* 2007). Enligt Kreplak *et al.* (2019) kan aktiviteten hos transposoner vara en del av förklaringen till ärtgenomets särdrag.

Jämförande genomik, på engelska Comparative genomics, är en analysmetod baserad på sekvenseringsdata där genomiska egenskaper, däribland DNA-sekvenser, gener eller genomiska markörer, hos organismer ställs mot varandra för att identifiera eventuella likheter och skillnader mellan arter eller accessioner (Touchman 2010). Att söka efter homologer hos närmare studerade arter inom Fabaceae kan vara ett sätt att identifiera gener av intresse hos ärta.

1.1.3. Rollen som proteingröda

Proteininnehållet i torra ärtfröer varierar mellan 13 och 30% beroende på sort (Tzitzikas *et al.* 2006), vilket motsvarar innehållet i trädgårdsböna (*Phaseolus vulgaris* L.) men är lägre än för sojaböna som innehåller upp till 40% protein (European Food Safety Authority 2014; Celmeli *et al.* 2018; Rizzo & Baroni 2018). Kommersiellt produceras i

huvudsak tre produkter baserade på ärtprotein, där ärtmjöl, ärtproteinkoncentrat och ärtproteinisolat listas från lägre till högre bearbetningsgrad och proteinandel (Day 2013).

En fördel med ärtor som proteinkälla är att den lämpar sig väl för ett nordeuropeiskt odlingsklimat (Leino & Nygård 2008). Den listas inte heller i Europeiska unionens lista över ingredienser med särskilda märknings- och informationskrav, till skillnad från soja som kan orsaka allvarliga allergiska reaktioner (Europeiska unionen 2011; European Food Safety Authority 2014). Livsmedelsverket rekommenderar trots detta att innehåll av ärtor eller ärtprotein i livsmedelsprodukter deklarerar då allergi mot ärtor förekommer (Sjögren Bolin 2021).

En utmaning i tillämpningen av ärtor och ärtprotein som livsmedel är dess smak. Flera studier visar att halten av i ärtor vanligt förekommande saponiner korrelerar med upplevd bitterhet (Price *et al.* 1985; Heng *et al.* 2006a). Smaken av saponiner i ärtor beskrivs även som sammandragande och metallisk (Price *et al.* 1985). Även andra ämnen, som exempelvis lipider, oxidationsprodukter av omättade fettsyror och vissa flyktiga ämnen, bidrar till en för ärtor karaktäristisk bismak (Murray *et al.* 1976; Heng 2005 se Roland *et al.* 2017; Gläser *et al.* 2020). Eftersom saponiner tenderar att binda till proteiner är denna ämnesgrupp av särskilt intresse inom växtförädlingsarbetet i syfte att förbättra proteinråvarans smakkvalitet och förutsättningar att användas inom livsmedelsindustrin (Heng *et al.* 2004; Gläser *et al.* 2020).

1.2. Syfte

Detta arbete är en del i forsknings- och växtförädlingsprojektet ”Ärtan – Garantin för framtidens gröna protein” inom SLU Grogrund, vilket har det långsiktiga syftet att ta fram nya sorters ärtor för att tillmötesgå efterfrågan på marknaden.

Syftet med arbetet är att undersöka vad som finns skrivet i litteraturen angående saponininnehåll i ärtor, då saponiner bidrar med en oönskad smak till livsmedel baserade på ärtor och ärtproteinprodukter. Vidare undersöks genom jämförande genomik förekomsten av homologer i ärtor till protein och gener involverade i soyasaponinsyntes i besläktade arter. Detta för att underlätta vidare forskning och växtförädling inom området.

1.3. Frågeställning

Vad är känt om bakomliggande genetiska mekanismer för soyasaponininnehåll i ärtor (*Pisum sativum* L.) och hur varierar saponinhalten beroende på genotyp och miljöfaktorer?

2. Metod

2.1. Litteratursökning

Flera databaser och sökmotorer har använts för litteratursökning, däribland Web of Science Core Collection, SLU Primo, Google och Google Scholar. I första hand har vetenskapligt granskade publikationer och originalverk använts, men även översiktsartiklar och rapporter från organisationer eller myndigheter förekommer som referenser.

Sökord som använts är exempelvis, men inte begränsade till, *Pisum*, "*Pisum sativum*", legumes, Fabaceae, saponin*, soyasaponin*, "DDMP saponin*", triterpen*, beta-amyrin, gene*, genom*, oxidosqualene, soyasapogenol, chromosaponin*.

2.2. Jämförande genomik

Analys av förekomsten av homologer i ärtor för gener och enzym av relevans för saponinsyntes i andra växtarter, primärt inom familjen ärtväxter (Fabaceae), har genomförts. Accessionsnummer (acc.nr.) för protein- och nukleotidsekvenser uppsöktes i databaser tillgängliga via National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov>), för att användas i sökverktyget Basic Local Alignment Search Tool (BLAST) tillgängligt via samma organisation. Sökningarna avgränsades till träffar inom arten *Pisum sativum* (taxid:3888).

För varje analyserad gen, eller sökobjekt, genomfördes två sökningar i NCBI Nucleotide Collection (NC) där bland annat gen-, mRNA- och proteinsekvenser finns tillgängliga. En med nukleotidsekvensen mot tillgängliga nukleotidsekvenser, Standard Nucleotide BLAST (blastn), och en med potentiella översättningar av proteinsekvensen mot nukleotidsekvenser, Translated BLAST (tblastn). Algoritmen för blastn optimerades för att tillåta sekvensträffar, eller subjekt, ner till sju baspar långa.

Vidare genomfördes en tblastn i databasen för Whole-Genome Shotgun (WGS) med sökobjekten, exklusive oxidosqualene-cyklaser. Presentationen av subjekt avgränsades till sekvenser inom projektet genomfört av The International Pea Genome Consortium (acc.nr. PRJEB31320, BioProject 542457, WGS-projekt CAADHX02), v1a, där sekvenserna är monterade i kopplingsgrupper motsvarande ärtas sju kromosomer (Kreplak *et al.* 2019). Information om var i genomet och på vilken kromosom subjekt i

resultatet återfinns var tillgänglig via en sökning i projektet via NCBI med BioProject-numret och nukleotidsekvensens accessionsnummer för subjektet i fråga.

För varje sökning redovisas, utöver subjektsekvensernas accessionsnummer, objektäckning (Query cover), det vill säga hur stor andel i procent av sökobjektets sekvens som täcks i subjektet, sekvensidentitet (Sequence identity), andelen nukleotider eller aminosyror inom subjektsekvensen som är identiska med de i objektsekvensen uttryckt i procent, samt E-värde (Expect value), ett mått på hur signifikant en jämförelse är. Ett E-värde nära 0,0 indikerar att resultatet med låg sannolikhet beror på slumpen (Kerfeld & Scott 2011).

För att analysera sekvensidentitet och objektäckning mellan två valda träffar kördes verktyget Align Sequences Protein BLAST via NCBI.

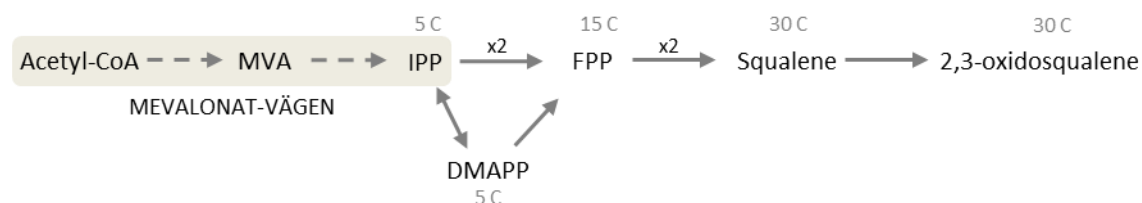
2.3. Avgränsningar

Vilken funktion saponiner har i växter och saponiners potentiella tillämpningar har i detta arbete inte undersökts. Inte heller hur eventuella eftersköldsbehandlingar kan påverka saponininnehållet diskuteras. Ingen analys av BLAST-resultaten har utförts i fråga om var i sekvensen objektäckningen ligger.

3. Resultat

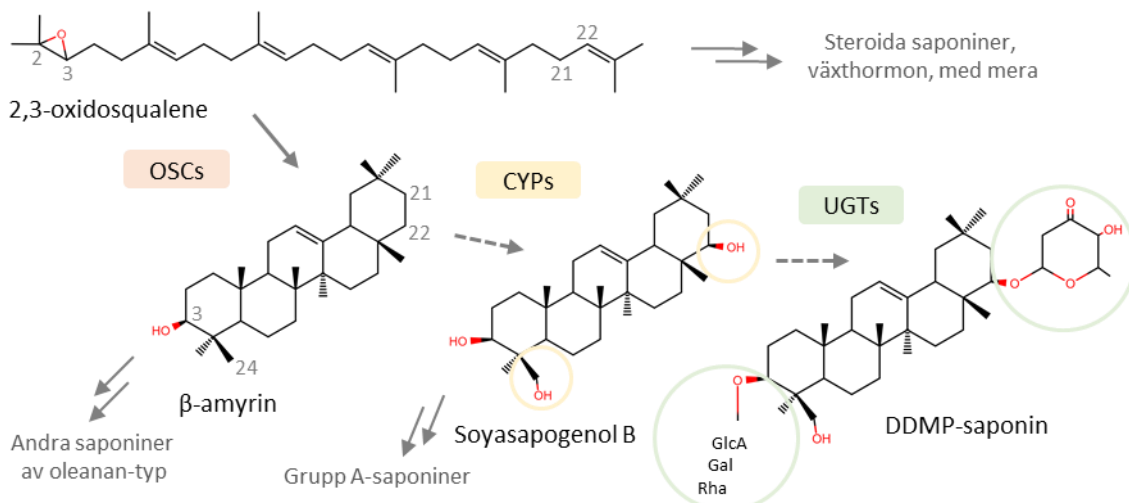
3.1. Saponiner – klassificering, kemi och biosyntes

Saponiner betraktas som sekundärmetaboliter och förekommer hos växter representerade i nära 100 familjer (Oleszek 2002). De tros bland annat fylla funktioner inom växternas försvar mot insekter och patogener då de har antimikrobiella- och insekticida egenskaper (Osbourn 2003; Sparg *et al.* 2004). De är icke-flyktiga glykosider bestående av en opolär aglykon, ur grekiskans nekande ”a” och ”glyky’s” som betyder söt, kopplad till en eller flera sockerkedjor (Hostettmann & Marston 1995; Taiz *et al.* 2015). Den kemiska strukturen, med en polär- och en opolär grupp ger saponiner och växter innehållande desamma tvållika egenskaper vilket historiskt haft betydelse för människan (Taiz *et al.* 2015). Tillämpningar förekommer även inom naturmedicin, läkemedelsindustrin och som färgämnen. De delas in i triterpenoida- eller steroida saponiner beroende på produkten efter cyklisering av föregångaren 2,3-oxidosqualene (Moses *et al.* 2014a).



Figur 3: Visualiserad och förenklad syntesväg för 2,3-oxidosqualene från acetyl-CoA (acetylkoenzym A). Mevalonat-vägen är en syntesväg i flera steg. Antalet kol i föreningarna indikeras i grå text. Katalyserande enzym exkluderade. MVA: Mevalonat; IPP: isopenetylpyrofosfat; DMAPP: dimetylallylpyrofosfat; FPP: farnesylpyrofosfat.

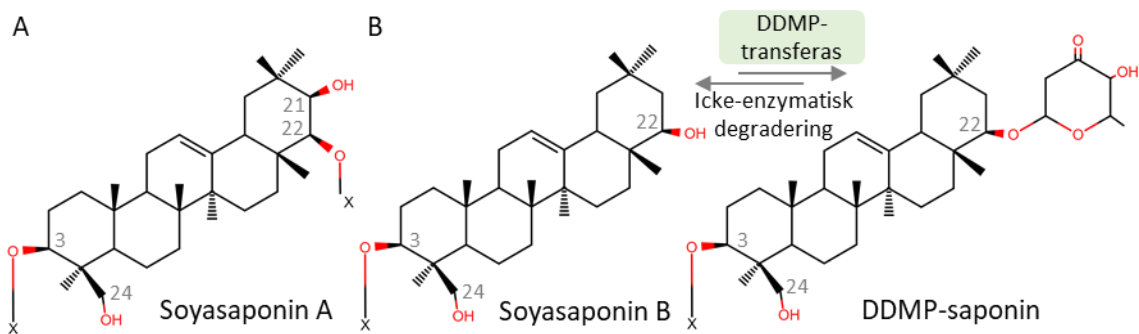
Det acykliska 2,3-oxidosqualene med 30 kolatomer erhålls genom en syntesväg i flera steg, från acetyl-koenzym A via bland andra isopenetylpyrofosfat och dess isomer dimetylallylpyrofosfat och är essentiell för flera organismgrupper, däribland djur och växter (Figur 3) (Moses *et al.* 2014a). Triterpener behåller 30 kolatomer i dess grundstruktur, medan grunden till steroida saponiner, som exempelvis cykloartenol och lanosterol, enbart har 27 kolatomer (Moses *et al.* 2014a). Cykloartenol är utöver saponiner även föregångare till bland andra kolesterol i djur och växthormon ur klassen brassinosteroider (Fujioka & Yokota 2003; Aboobucker & Suza 2019). Saponiner av oleanan-typ sägs vara de som erhålls ur den pentacykliska grundstrukturen β -amyrin (3 β -hydroxyolean-12-ene) och innefattar bland andra glycyrrhizin i lakrits (*Glycyrrhiza* spp.), avenacin i havre (*Avena* spp.) och soyasaponiner (Thimmappa *et al.* 2014; Chung *et al.* 2020).



Figur 4: I soja föreslagen syntesväg och skelettformler för ämnena 2,3-oxidosqualene, β -amyrin, soyasapogenol B och 2,3-dihydro-2,5-dihydroxy-6-metyl-4H-pyran-4-one (DDMP)-saponin. Soyasaponin Bb skiljer från DDMP-saponin enbart avseende DDMP-delen som saknas hos den förstnämnda. Oxidosqualene-cyklaser (OSCs), cytokrom P450-beroende monooxygenaser (CYPs) och uridindifosfat-beroende glykosyltransferaser (UGTs) sägs vara enzymgrupper som är vanligen involverade i syntesen. Numrering av kolen i grundstrukturen indikeras med grå siffror. Sockerarterna vid kol 3 i soyasaponinerna i ärtä är glukuronsyra (GlcA), galaktos (Gal), och rhamnos (Rha). Streckade pilar indikerar att flera enzym är involverade där intermediärer inte är gestaltade, dubbla pilar indikerar exempel på andra metaboliter som kan erhållas ur samma föregångare.

Soyasaponiner är triterpenoida saponiner av oleanan-typ och är vanligt förekommande i växter inom familjen ärtväxter (Krishnamurthy *et al.* 2020). De erhålls ur β -amyrin som genomgår hydroxylering och bildar soyasapogenol, följt av glykosylering och andra modifieringar vilka resulterar i bildandet av soyasaponiner (Figur 4) (Moses *et al.* 2014a). Vanligen kategoriseras de till grupp A- eller 2,3-dihydro-2,5-dihydroxy-6-metyl-4H-pyran-4-one (DDMP)-saponiner (Sundaramoorthy *et al.* 2019), men även grupp B-saponiner och andra mindre vanliga grupper har påträffats (Kudou *et al.* 1993; Daveby *et al.* 1998).

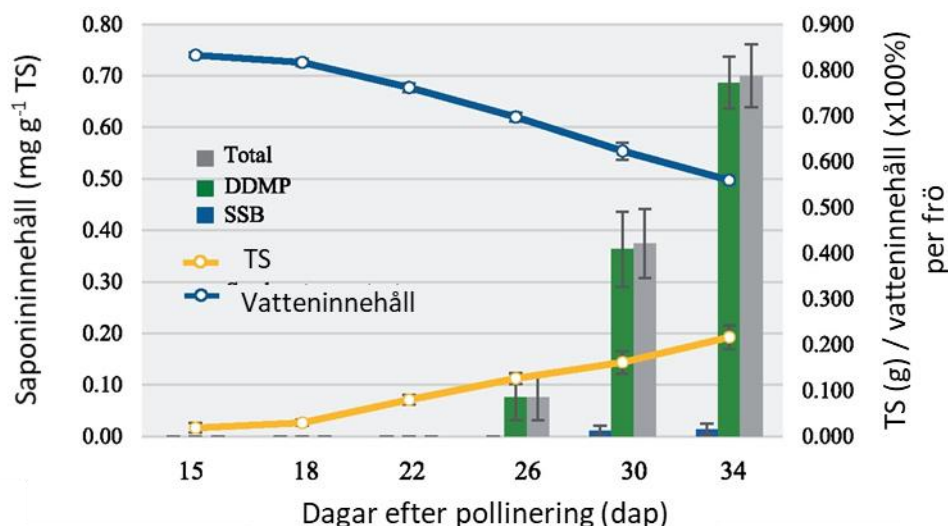
Grupp B-saponiner erhålls ur aglykonen soyasapogenol B (3 β ,22 β ,24-trihydroxyolean-12-ene), som bildas genom hydroxylering av kol 22 (C-22) och C-24 hos β -amyrin, och har en ensam sockerkedja bunden till C-3 (Figur 5B) (Sayama *et al.* 2012). Grupp A-saponiner särskiljer sig genom att ha ytterligare en sockerkedja bunden till C-22 (Figur 5A), och erhålls ur soyasapogenol A (3 β ,21 β ,22 β ,24-tetrahydroxyolean-12-ene) som även genomgår en hydroxylering av C-21 (Sawai & Saito 2011; Sayama *et al.* 2012). DDMP-saponiner särskiljs från grupp B-saponiner genom DDMP-delen vid C-22 (Figur 5B) (Sawai & Saito 2011).



Figur 5: Skelettformel för (A) soyasaponin A och (B) soyasaponin B och DDMP-saponin, samt omvandlingen däremellan. Grå siffror indikerar kolens numrering utifrån den ursprungliga ordningen i det acykliska 2,3-oxidosqualene. DDMP: 2,3-dihydro-2,5-dihydroxy-6-methyl-4H-pyran-4-one; X: sockerkedjor.

Till skillnad från i soja och flera andra baljväxter som ackumulerar både saponiner ur grupp A och B, har i ärtä enbart DDMP-saponinet soyasaponin β g och grupp B-saponinet soyasaponin Bb påvisats (figur 5) (Heng *et al.* 2006a). Flera synonymer förekommer i litteraturen, och i arbetet används hädanefter DDMP-saponin respektive soyasaponin Bb för dessa ämnen. Den förstnämnda upplevs som bitter och kan hydrolyseras till soyasaponin Bb och maltol vid exempelvis extraktion, upphettning över 40 grader eller lagring (Figur 5B) (Kudou *et al.* 1993; Daveby *et al.* 1998; Heng *et al.* 2006a; Heng *et al.* 2006b). Vedertaget är att grupp A- och DDMP-saponiner ackumuleras genom biosyntes, medan grupp B-saponiner är rester från syntes eller hydrolys av DDMP-saponin (Sundaramoorthy *et al.* 2019; Vernoud *et al.* 2021).

Syntes och ackumulering av soyasaponiner i ärtväxter sker främst i rötter, noder och fröer, men även till viss del i blad och stam (Oleszek & Stochmal 2002; Huhman *et al.* 2005; Faizal & Geelen 2013; Tsuno *et al.* 2018; Vernoud *et al.* 2021). Studier i både soja och ärtä visar att fröer innehåller betydande mängder soyasaponiner, men att ackumuleringen begränsas till hypokotyl, kotyledon och embryo medan fröskalerna inte innehåller några detekterbara mängder (Kitagawa *et al.* 1988; Shiraiwa *et al.* 1991a; Daveby *et al.* 1997; Vernoud *et al.* 2021). Vidare finns studier som visar att soyasaponiner i ärtä ackumuleras i fröet under mognadsfasen för att kulminera vid mognad, 34 dagar efter pollinering (dap) (Figur 6) (Vernoud *et al.* 2021), och att liknande ackumuleringsmönster kan observeras för β -amyrin, föregångaren till soyasaponiner (Baisted 1971). Soyasaponin och β -amyrin ackumuleras även i fröet under imbibering och groning (Baisted 1971; Vernoud *et al.* 2021).



Figur 6: Saponininnehåll i ärtfröer av sorten 'Caméor' under fröutveckling. Saponininnehåll på den vänstra axeln, samt torrsbstans (TS) och vattenhalt per frö på den högra axeln. Felstaplar representerar standardavvikelse från medelvärde av tre biologiska replikat. DDMP: 2,3-dihydro-2,5-dihydroxy-6-methyl-4H-pyran-4-one-saponin; SSB: soyasaponin Bb. Figur erhållen från Vernoud et al. (2021) och översatt från engelska med tillstånd från Oxford University Press.

I syntesen av triterpener anses tre huvudsakliga kategorier av enzymer vara avgörande (Figur 4), oxidosqualene-cyklaser (OSCs), cytokrom P450-beroende monooxygenaser (CYPs) och uridindifosfat-beroende glycosyltransferaser (UGTs), även om andra enzym, i synnerhet transferaser, kan vara involverade (Sawai & Saito 2011; Moses *et al.* 2014a). Eftersom en helgenomsekvensering av ärt publicerades 2019 har möjligheten att utföra jämförande genomik innan dess varit begränsad till proteinsekvenser och kortare sekvenserade DNA-fragment (Kreplak *et al.* 2019). I dagsläget finns enbart två identifierade gener vars funktion har kopplats till syntesen av soyasaponiner i ärt, vilka utgörs av OSCs (Morita *et al.* 1997; Morita *et al.* 2000; Vernoud *et al.* 2021).

3.2. Enzym kopplade till soyasaponinsyntes i ärtväxter

I följande tre avsnitt presenteras de grupper av enzym som har kopplats till syntesen av soyasaponiner. Funktionen för specifika enzym i ärt och andra baljväxter redogörs för, och resultat från sökningar i BLAST och jämförande genomik sammanfattas.

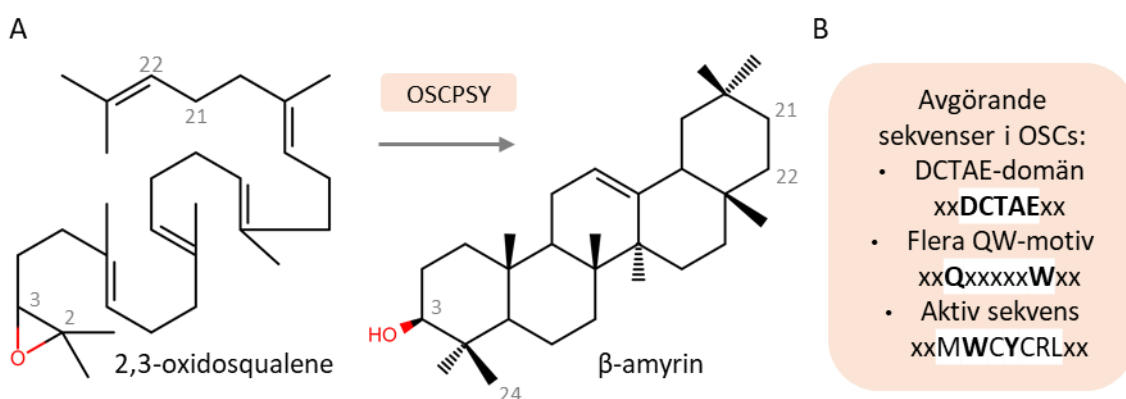
I Bilagor finns en sammanställning av växtenzym vars funktion identifierats som involverade i syntesen av soyasaponiner eller ämnen som soyasaponiner potentiellt kan erhållas ur (Tabell B1). Funktionen av enzymer kopplade till triterpensyntes är främst utvärderad genom homologt uttryck av cDNA i transgena celler av jästsvamp (*Saccharomyces cerevisiae*) (Thimmappa *et al.* 2014), men även andra metoder som inducerad mutagenes eller homologt genuttryck i modellväxter förekommer.

Sökningar i BLAST har generellt avgränsats till protein och gener relevanta för soyasaponinsyntesen i arter inom familjen ärtväxter (Fabaceae). Sökobjekt samt resultat från sökningarna finns listade i Tabell B2 och resultat i BLAST hänvisar till denna eller till löpande tabeller om ingen annan källa anges. Alla E-värden för sökträffar presenterade i resultatet är mindre än 3e-04.

3.2.1. Oxidosqualene-cyklaser

Enligt Thimmappa *et al.* (2014) är cykliseringen av 2,3-oxidosqualene en av de mest komplexa enzymatiska reaktionerna i metabolismen av triterpener hos växter. Identifierade OSCs är i många fall multifunktionella, men i hög grad specialiserade enzymer är också vanligt förekommande (Thimmappa *et al.* 2014). De OSCs av relevans för soyasaponinsyntesen är β -amyrin-syntaser eller multifunktionella enzymer som delvis producerar β -amyrin (Figur 7A) (Moses *et al.* 2014b).

I OSCs, som β -amyrin-syntaser men även cykloartenol- och lanosterol-syntaser, finns flera konserverade aminosyrasekvenser vilka är avgörande för enzymernas funktion och produktspecificitet (Figur 7B).



Figur 7: Cyklisering av 2,3-oxidosqualene till β -amyrin. (A) Skelettformler för substrat och produkt, här katalyserat av β -amyrin-syntaset OSCPSY (acc.nr. AB034802) i ärtä. Grå siffror indikerar kolens numrering utifrån den ursprungliga ordningen i det acykliska 2,3-oxidosqualene. (B) Oxidosqualene-cyklasers (OSCs) funktion karaktäriseras av ett antal aminosyrasekvenser.

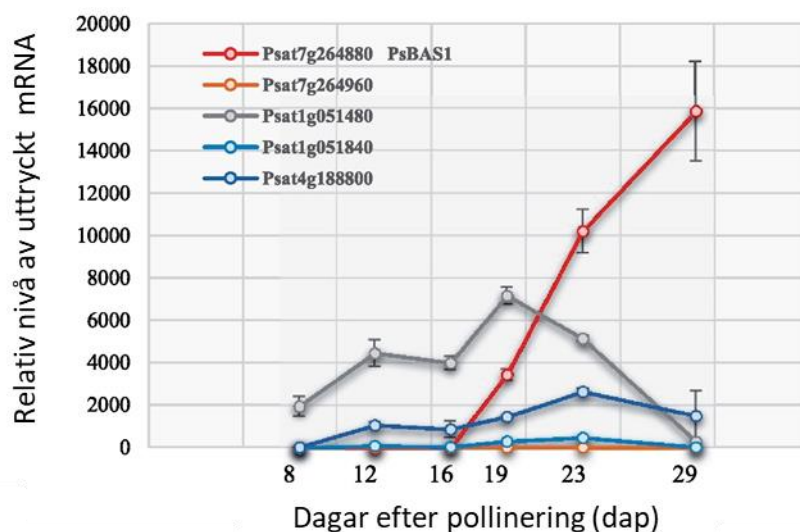
DCTAE-domänen, med aminosyrasekvensen asparaginsyra-cystein-treonin-alanin-glutaminsyra (DCTAE), är involverad i inbindning och protonation av 2,3-oxidosqualene, och därmed i initieringen av cykliseringsreaktionen (Wang *et al.* 2011; Ito *et al.* 2013). QW-motiv är sekvenser i enzymet om sju aminosyror, som börjar med glutamin (Q) och slutar med tryptofan (W), vilka konformerar det acykliska substratet till dess cykliska struktur (Poralla *et al.* 1994). I β -amyrin-syntaser i soja konstateras att glutamin och tryptofan är konserverade medan de fem aminosyrorna däremellan varierar (Ali *et al.* 2016). För polycykliska produkter krävs återkommande QW-motiv, och

upprepad förekomst av QW-motiv anses vara unikt för OSCs (Poralla *et al.* 1994). I OSCs har även en konserverad aktiv sekvens identifierats, i vilken tyrosin (Y) är avgörande för bildandet av den pentacykliska strukturen i allmänhet och tryptofan (W) för produktdifferentiering till β -amyrin i synnerhet (Kushiro *et al.* 2000a; Vernoud *et al.* 2021).

Tre enzym i ärtor har identifierats vars cDNA i transgena jästceller uttrycker OSCs (Morita *et al.* 1997; 2000). Den tidigare studien beskriver ett cykloartenol-syntas, CASPEA (acc.nr. D89619), vilken identifierats genom att klonas cDNA med hjälp av PCR-primers baserade på konserverade homologa sekvenser från kända cyklaser hos andra växt- och djurarter. CASPEA katalyserar bildandet av cykloartenol som är föregångare till flera steroida saponiner (Corey *et al.* 1993; Hostettmann & Marston 1995; Morita *et al.* 1997), vilka inte har någon känd relevans för den oönskade bismaken i ärtfröer och ärtprotein (Roland *et al.* 2017).

I den senare studien kompletteras metoden med PCR-primers baserade på nukleotidsekvenser unika för β -amyrin-syntaser hos andra växtarter, i syfte att förhindra amplifiering av cykloartenol-syntaser (Morita *et al.* 2000). Därigenom identifieras OSCPSY och OSCPSM (acc.nr. AB034802 respektive AB034803), vars funktion visas vara ett β -amyrin-syntas respektive ett multifunktionellt α - och β -amyrin-syntas (Morita *et al.* 2000).

Via en sökning med proteinsekvensen för OSCPSY i BLAST mot ärtas helgenomsekvens, v1a, identifieras flera i hög grad homologa nukleotidsekvenser (Vernoud *et al.* 2021). Fem av dessa innehåller en DCTAE-domän och flera QW-motiv. Sekvensen för *PsBAS1* är identisk med den översatta proteinsekvensen från OSCPSY och återfinns på kromosom 7, Psat1g051480 med sekvensen från cykloartenol-syntaset CASPEA och Psat4g188800 är till 98% identisk med sekvensen från det multifunktionella enzymet OSCPSM (Vernoud *et al.* 2021). Vidare är Psat1g051480 i hög grad homolog med *PsBAS1* men saknar delar av sekvensen, däribland den aktiva sekvensen som hos övriga sekvenser finns konserverad. Extrahering av mRNA ur fröer vid sex tillfällen under fröutvecklingen visar uttrycksmönstret för dessa sekvenser (Figur 8) (Vernoud *et al.* 2021).



Figur 8: Relativ nivå av uttryckt mRNA för flera förmodade oxidosqualene-cyklaser (OSCs) under fröutveckling i ärtsorten 'Caméor'. PsBAS1 är identisk med β -amyrin-syntaset OSCPSY och kan kopplas till syntesen av soyasaponiner i ärt. Psat1g051480 är identisk med cykloartenol-syntaset CASPEA, Psat4g188800 med det multifunktionella enzymet OSCPSM. Felstaplar representerar standardavvikelse från medelvärde av tre biologiska replikat. Figur erhållen från Vernoud et al. (2021) och översatt från engelska med tillstånd från Oxford University Press.

I Figur 8 korrelerar 8 dap med embryogenes, 12 dap med övergången från embryogenes till fyllnad av fröet, 16, 19 och 23 dap med fyllnad av fröet och 29 dap med begynnande fröabskission (Vernoud *et al.* 2021). Uttrycket för *PsBAS1* ökar gradvis under mognad från och med 18 dap och når det maximala värdet vid sista extraktionstillfället, vilket överensstämmer med ackumuleringen av DDMP-saponin i fröerna (Figur 6) (Vernoud *et al.* 2021). Studien visar även att mRNA från *PsBAS1* uttrycks i fröerna efter imbibering.

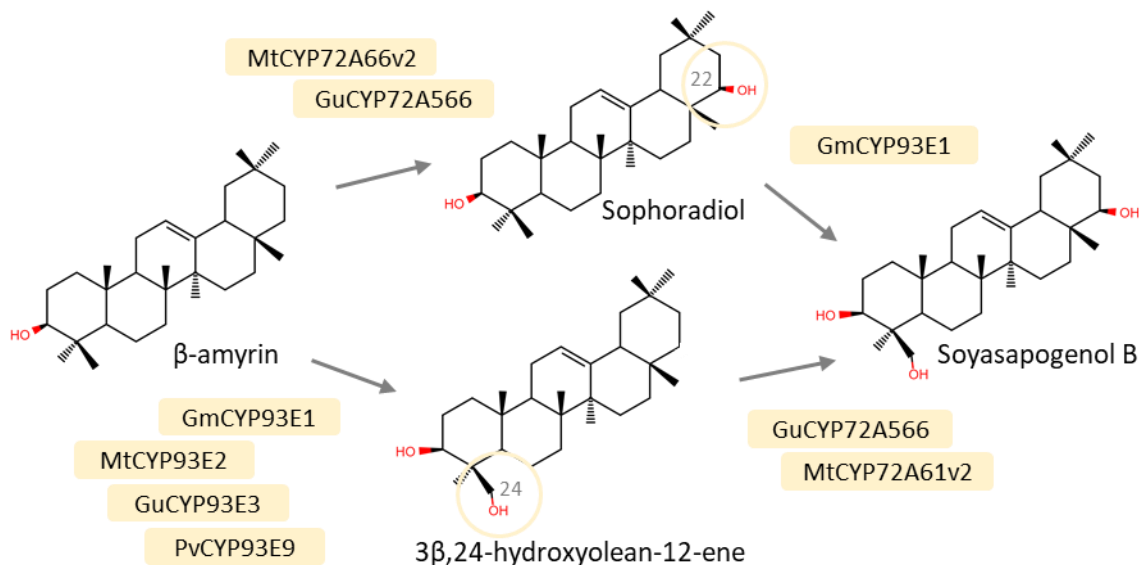
Genom inducerad mutagenes, där punktmutationer i *BAS1* som direkt eller indirekt resulterade i tidiga stopp-kodon och förlusten av flera för funktionen avgörande aminosyrasekvenser i muterade enzym, konstateras att individer utan funktionellt OSCPSY inte ackumulerar någon detekterbar mängd av någon i vildtypen förekommande soyasaponiner, varken i fröer eller i andra organ (Vernoud *et al.* 2021).

Resultaten av sökningar i BLAST med OSCs från arter inom ärtordningen visar att samtliga analyserade enzym ger träffar på OSCPSY, OSCPSM och CASPEA i ärt (Tabell B2). För OSCPSY och OSCPSM är sekvensidentiteten över 69% för samtliga jämförelser i blastn, medan homologin är något lägre för cykloartenol-syntaset CASPEA. För sökningar med översatta proteinsekvenser (tblastn) har samtliga jämförelser en sekvensidentitet mellan 56 och 59% för CASPEA respektive 77 och 95% för OSCPSY och OSCPSM.

3.2.2. Cytokrom P450-beroende monooxygenaser

CYPs är en stor grupp enzymer med funktioner inom växters primära- och sekundära metabolism (Thimmappa *et al.* 2014; Cook *et al.* 2016), och representerar runt 1% av alla proteinkodande gener hos växter (Nelson & Werck-Reichhart 2011). Karaktäristiskt för enzymgruppen är bland annat en hydrofobiskt järninnehållande del, flera spiralformade helix-segment och konserverade aminosyrasekvenser vid både de N- och C-terminala ändarna (Moses *et al.* 2014b; Cook *et al.* 2016). Enzym inom gruppen delar vanligtvis låg sekvensidentitet (Moses *et al.* 2014b). Många CYPs finns lokaliserade i lumen i växtcellernas endoplasmatiska retikulum, men kan även påträffas i andra membranstrukturer (Ro *et al.* 2001). Enzymerna delas upp i familjer vidare indelade i underfamiljer, baserade på dess sekvenslikhet (Cook *et al.* 2016). Fylogenetisk analys tyder på att alla CYPs i växter härstammar från en ursprunglig gen (Nelson 1999).

I syntesen av soyasaponiner är CYPs ansvariga för hydroxyleringen av β -amyrin till soyasapogenoler, vilken förmodas ske i två steg med flertalet enzym involverade (Sawai & Saito 2011; Thimmappa *et al.* 2014). Beroende på om C-22 eller C-24 hydroxyleras först går syntesen via sophoradiol (3 β ,22-hydroxyolean-12-ene) respektive 3 β ,24-hydroxyolean-12-ene (Figur 9) (Shibuya *et al.* 2006).



Figur 9: Föreslagen syntesväg av hydroxylering från β -amyrin till soyasapogenol B, via sophoradiol (3 β ,22-hydroxyolean-12-ene) eller 3 β ,24-hydroxyolean-12-ene. Gula boxar indikerar enzym som funktionellt visats katalysera reaktionen i olika växter inom växtfamiljen Fabaceae. Grå siffror indikerar kolens numrering utifrån den ursprungliga ordningen i det acykliska 2,3-oxidosqualene. Gm: *Glycine max*; Gu: *Glycyrrhiza uralensis*; Mt: *Medicago truncatula*; Pv: *Phaseolus vulgaris*.

Ur underfamiljen CYP93E finns fyra identifierade enzym hos växter inom Fabaceae vars funktion kan kopplas till hydroxylering av C-24 hos β -amyrin (Tabell B1; Figur 9) (Shibuya *et al.* 2006; Seki *et al.* 2008; Fukushima *et al.* 2011; Moses *et al.* 2014b). Av

dessa har CYP93E1 i soja även konstaterats hydroxylera sophoradiol vid C-24 och en funktionsanalys av enzymet visar att hydroxylaset är substratspecifikt för 3 β -hydroxyolean-12-ene-strukturen som återfinns i β -amyrin och sophoradiol (Shibuya *et al.* 2006). CYP93E1 påstås vidare vara det första identifierande hydroxylaset av triterpener i växter.

Resultat från BLAST påvisar att samtliga enzym ur underfamiljen CYP93E i arter använda som sökobjekt i denna studie påvisar hög homologi mot enzymet CYP93E8 (acc.nr. KF906539.1) i ärt, både för nukleotid- och proteinsekvenserna (Tabell 1). I en studie på enzym tillhörande underfamiljen CYP93E av Moses *et al.* (2014b) föreslås genom datasimulering att CYP93E8 i ärt är en potentiell katalysator i omvandlingen av β -amyrin till intermediären 3 β ,24-hydroxyolean-12-ene. Vidare påvisas vid karaktärisering av funktion i transgena jäststammar låg katalytisk effekt med β -amyrin som substrat (Moses *et al.* 2014b).

Tabell 1: Urval av träffar från sökning i BLAST med cytokrom P450-beroende C-24-hydroxylas ur underfamiljen CYP93E.

Metod	Sökobjekt	Subjekt			
	Enzym	OT %	ID %	Acc.nr.	Information
Tblast NC	GmCYP93E1	99	80,4	KF906539.1	CYP93E8. Låg katalytisk effekt (Moses <i>et al.</i> 2014b)
	MtCYP93E2	99	90,4		
	GuCYP93E3	99	87,1		
	PvCYP93E9	99	77,7		
Tblast WGS	GmCYP93E1	99	80,1	CAADHX020004390.1	Kromosom 3
	MtCYP93E2	99	90,1		
	GuCYP93E3	99	80,2		
	PvCYP93E9	99	70,1		

Acc.nr.: accessionsnummer; BR: brassinosteroider; Gm: *Glycine max*; Gu: *Glycyrrhiza uralensis*; ID: sekvensidentitet; Mt: *Medicago truncatula*; NC: NCBI Nucleotide Collection; OT: objektäckning; Pv: *Phaseolus vulgaris*; WGS: whole-genome shotgun.

De analyserade proteinsekvenserna inom underfamiljen CYP93E ger även träffar i WGS-databasen på en sekvens som kan kopplas till kromosom 3 (acc.nr. CAADHX020004390.1) i ärt (Tabell 1), samt på ett odefinierat CYP (acc.nr. AF218296.1) (Tabell B2). Detta CYP-enzym kan kopplas till brassinosteroidsyntes och etiolering i groende ärt (Kang *et al.* 2001).

Enzymen CYP72A566 i kinesisk lakrits (*Glycyrrhiza uralensis* L.) och CYP72A61v2 i tornlusern har katalytisk effekt av hydroxyleringen vid C-22 hos β -amyrin när uttryckta i transgena jästceller (Figur 9) (Fukushima *et al.* 2013; Tamura *et al.* 2018), och har en inbördes sekvensidentitet på 85% (Tamura *et al.* 2018). Enzymet CYP72A66v2 i tornlusern kan katalysera hydroxylering vid C-22, men även andra reaktionsprodukter bildas (Fanani *et al.* 2019). CYP72A61v2 saknar förmågan att använda β -amyrin som

substrat (Seki *et al.* 2011). Vid försök att tysta uttrycket av *CYP72A61v2* med riktad mutagenes erhöles inga livsdugliga embryon (Confalonieri *et al.* 2021).

Sökning i BLAST med proteinsekvenserna från C-22-hydroxylasen påvisar hög objektäckning mot delar ur sekvenser kopplade till kromosom 7 och 1 (acc.nr. CAADHX020007147.1 respektive CAADHX020023846.1) i ärtä (Tabell 2). Genom Tblast i NC-databaser påvisas viss homologi även mot enzymet *CYP72A698* (acc.nr. MK534536.1) i ärtä. Ingen funktion av *CYP72A698* i ärtä finns beskriven.

Tabell 2: Urval av träffar från sökning i BLAST med cytokrom P450-beroende hydroxylas ur underfamiljen CYP72A.

Metod	Sökobjekt	Subjekt			
	Enzym	OT %	ID %	Acc.nr.	Information
Tblast NC	GuCYP72A566	96	49,5	MK534536.1	CYP72A698. Ingen påvisad katalytisk effekt (Fanani <i>et al.</i> 2019)
	MtCYP72A61v2	92	51,3		
	MtCYP72A66v2	91	57,1		
Tblast WGS	GuCYP72A566	97	88,9	CAADHX020007147.1	Kromosom 7
	MtCYP72A61v2	100	93,6		
	MtCYP72A66v2	90	61,7		
	GuCYP72A566	97	66,0	CAADHX020023846.1	Kromosom 1
	MtCYP72A61v2	96	63,8		
	MtCYP72A66v2	91	89,5		

Acc.nr.: accessionsnummer; Gm: *Glycine max*; Gu: *Glycyrrhiza uralensis*; ID: sekvensidentitet; Mt: *Medicago truncatula*; NC: NCBI Nucleotide Collection; OT: objektäckning; WGS: whole-genome shotgun.

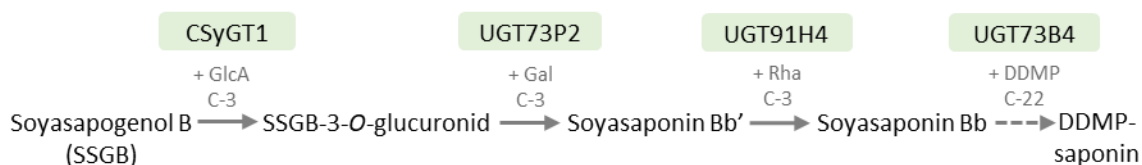
Proteinsekvenserna ur underfamiljen CYP72A ger även träff på CYP734A11 (acc.nr. AB218764.1) i ärtä (Tabell B2), en homolog till brassinosteroid-regulatorproteinet CYP734A1 från *Arabidopsis thaliana* (Nomura *et al.* 2007). I ärtä uttrycks mRNA från CYP734A11 i hög grad i omoget frö och under imbibering, och detekteras under frömognad och moget frö (Nomura *et al.* 2007).

3.2.3. Glykosyltransferaser

Glykosyleringsteget i saponininsynthesen innebär att socker binder till aglykonen, vilket leder till ökad polaritet och bidrar till de egenskaper som karaktäriserar saponiner (Thimmappa *et al.* 2014). Enzymer involverade är bland andra glykosyltransferaser (GTs), i synnerhet uridindifosfat (UDP)-beroende glykosyltransferaser (UGTs), men även andra typer av transferas kan bidra till modifieringen (Moses *et al.* 2014a). GTs och dess reaktioner antas äga rum i cytosolen, eventuellt i association med proteinkomplex i membranstrukturer eller med transmembranprotein (Bowles *et al.* 2006).

UGTs kan vara mer eller mindre specialiserade, men många tycks vara regioselektiva (Vogt & Jones 2000), vilket innebär att glykosylering företrädesvis sker vid en specifik acceptor hos substratet av flera möjliga (Nationalencyklopedin u.å.-b). Inbindning till substratet som genomgår glykosylering tros ske vid regioner i den N-terminala änden av enzymet (Yonekura-Sakakibara & Hanada 2011). Alla GTs har vid den C-terminala änden en konserverad sekvens om 44 aminosyror, kallad PSPG (Plant Secondary Product Glycosyltransferase)-box, som tros vara involverad i inbindningen till UDP (Hundle *et al.* 1992; Yonekura-Sakakibara & Hanada 2011). Sockret erhålls vanligen ur UDP-bundna monosackarider (Thimmappa *et al.* 2014), och variationer i PSPG-boxsekvensen är kopplade till sockerdonatorspecificitet (Kubo *et al.* 2004).

Soyasaponin Bb och DDMP-saponin hos ärta har en sockerkedja bunden till C-3, vars motsvarigheter i soja adderas till soyasapogenol B i flera steg där enzym som katalyserar samtliga reaktioner finns identifierade (Figur 10) (Chung *et al.* 2020; Krishnamurthy *et al.* 2020). Donatorsubstraten som nyttjas är UDP-glucuronsyra, -galaktos respektive -rhamnos (Shibuya *et al.* 2010; Chung *et al.* 2020). För inbindningen av DDMP-delen antas UDP-DDMP vara donator (Sundaramoorthy *et al.* 2019). Liksom för CYPs delas UGTs in i familjer och underfamiljer utifrån aminosyrasekvensernas homologi (Ross *et al.* 2001).



Figur 10: Föreslagen syntesväg i soja (*Glycine max*) från soyasapogenol B till DDMP-saponin. Katalyserande enzym i gröna boxar. I figuren katalyserar UGT73B4 inbindningen av DDMP-delen till soyasaponin Bb, men kan även nyttja soyasapogenol B som acceptorsubstrat. GlcA: glucuronsyra; Gal: galaktos; Rha: rhamnos; DDMP: 2,3-dihydro-2,5-dihydroxy-6-methyl-4H-pyran-4-one.

Det senaste enzymet att identifieras i syntesen av soyasaponiner i soja är ett cellulosasyntas-erhållet GT, GmCSyGT1, vilket katalyserar inbindningen av glucuronsyra, en glukosliknande karboxylsyra, vid C-3 hos soyasapogenol B (Figur 10) (Chung *et al.* 2020). Vidare antas CSyGTs vara transmembranenzymer.

Varken enzymen UGT73K1 i tornlusern eller UGT73P12 i kinesisk lakrits har någon dokumenterad funktion i syntesen av soyasaponiner relevanta för ärta (Tabell B1) (Achnine *et al.* 2005; Nomura *et al.* 2019). UGT73K1 kan nyttja soyasapogenol B som acceptor, men är specialiserad att använda UDP-glukos som sockerdonator (Achnine *et al.* 2005). Enzymen UGT73P2 och UGT91H4 i soja kan däremot vara av intresse för syntesen av i ärta förekommande soyasaponiner då de katalyserar inbindningen av galaktos respektive rhamnos till sockerkedjan vid C-3 (Figur 10) (Shibuya *et al.* 2010).

Enzymet UGT73B4 i soja har identifierats katalysera inbindningen av DDMP-delen till C-22 och kan nyttja både soyasaponin Bb och aglykonen soyasapogenol B som substrat (Figur 10) (Sundaramoorthy *et al.* 2019). Genom inducerad mutagenes påvisas en signifikant minskning av ackumulerat DDMP-saponin i fröer hos homozygota mutanter jämfört med hos vildtypen, medan halterna i övriga organ är opåverkade. Att UGT73B4 är organspecifikt bekräftas även genom att studera uttrycket av mRNA i olika organ och under olika utvecklingsstadier som visar att genen uttrycks i högre grad i fröer än i andra organ (Sundaramoorthy *et al.* 2019).

Tabell 3: Urval av träffar från sökning i BLAST med glykosyltransferaser involverade i syntesen av soyasaponin Bb och DDMP-saponin i soja (*Glycine max*).

Metod	Sököbjekt	Subjekt			
	Enzym	OT %	ID %	Acc.nr.	Information
Tblast WGS	GmCSyGT1	100	72,6	CAADHX020000520.1	Kromosom 1
	UGT73P2	97	78,9	CAADHX020024174.1	Inte monterad
	UGT91H4	98	74,6	CAADHX020006075.1	Kromosom 1
	UGT73B4	97	63,1	CAADHX020012292.1	Kromosom 7

Acc.nr.: accessionsnummer; ID: sekvensidentitet; OT: objektäckning; WGS: whole-genome shotgun.

I BLAST-sökningar med GTs i soja mot helgenomsekvenserna i ärta visas samtliga dela en sekvensidentitet över 63% med en subjektsekvens vardera (Tabell 3). Subjekten återkommer som sökträffar med lägre sekvensidentitet mot jämförda enzymer ur samma underfamilj (Tabell B2). Sekvensen kopplad till kromosom 7 (acc.nr. CAADHX020012292.1) ger en träff med hög objektäckning mot enzymet UGT73K1 i tornlusern och med en sekvensidentitet på 84%.

BLAST-sökning i NC-databaser med nukleotidsekvenserna för enzymer GmCSyGT1, UGT73B4 och UGT91H4 i soja ger inga träffar i ärta, och resterande jämförda GTs ger träffar med låg objektäckning (Tabell B2). Sökningar med proteinsekvenserna i NC-databaser ger träffar på abskissat-beta-glucosyltransferaser (acc.nr. MF034881.1, MF034884.1 och KX518352.1) i ärta, med sekvensidentitet mellan 36 och 48% mot samtliga analyserade GTs (Tabell B2). För dessa gener i ärta finns ingen beskriven funktion publicerad.

3.3. Variation i saponinhalt

Saponinhalten i fröer av ärta har vid flera tillfällen konstaterats variera (Daveby *et al.* 1997; Heng *et al.* 2006a; Vernoud *et al.* 2021). Vad som är känt kring vilka faktorer som styr variationen samt gener av intresse presenteras i kommande avsnitt, och resultat från sökning i BLAST sammanfattas.

3.3.1. Faktorer som styr variationen

Hur saponininnehållet varierar mellan arter med olika genetisk bakgrund, fenotyp och slutanvändning har undersökts vid flertalet tillfällen. Analys av saponinhalten i minst 20 accessioner av *Pisum sativum* visar att innehållet i fröer varierar mellan 0,82 – 2,5 g kg⁻¹ torrsvikt, och att inget signifikant samband kan uttydas mellan variationen och vilket år eller på vilken plats fröerna odlats (Daveby *et al.* 1997). Senare studier använder varsammare metoder för extrahering där det totala innehållet i 16 respektive 11 olika accessioner av *Pisum* spp. varierar mellan 0,7 – 1,9 g kg⁻¹ (Heng *et al.* 2006a) respektive 0,36 – 1,2 g kg⁻¹ (Vernoud *et al.* 2021). De varsammare metoderna visar vidare på lägre maximala värden samt att DDMP-saponin utgör merparten av det totala saponininnehållet i samtliga analyserade fröer, till skillnad från Daveby *et al.* (1997) som enbart påvisar förekomst av soyasaponin Bb.

Sambandet mellan saponininnehåll och fenotyp för 54 accessioner tillhörande *Pisum sativum* har analyserats genom att gruppera accessionerna i följande tre kategorier, vitblommig foderärt, färgade fröer respektive skrynkliga fröer (Bastianelli *et al.* 1998). Resultatet påvisar ingen signifikant skillnad mellan kategorierna, vilket även bekräftas av Daveby *et al.* (1997). Ett samband där sorter med gula fröer tenderar att ha lägre saponininnehåll än sorter med gröna eller bruna fröer har senare påvisats (Heng *et al.* 2006a). Vernoud *et al.* (2021) visar att av 11 analyserade accessioner är de två sorterna med lägst saponininnehåll de enda i jämförelsen med lila blommor, men ingen information om huruvida korrelationen är signifikant kan erhållas.

Studier om saponininnehåll i soja och blålupin (*Lupinus angustifolius* L.) bekräftar att genotypen är avgörande för variationen hos dessa arter (Shiraiwa *et al.* 1991b; Ruiz *et al.* 1995), men att även miljöfaktorer konstateras ha en viss inverkan för innehållet i blålupin (Ruiz *et al.* 1995).

Det finns indikationer som tyder på att både biotisk- och abiotisk stress kan trigga syntesen av soyasaponiner i växter. Hur symbios med de kvävefixerande bakterierna inom *Rhizobium* spp. påverkar toleransen mot biotisk stress och fröskördens kvalitet har undersökts i två sorter av ärt (Ranjbar Sistani *et al.* 2017). Studien visar att ackumuleringen av soyasapogenol i fröerna ökar vid infektion av en patogen svamp. Det påvisas också att ackumulerat soyasapogenol är högre i fröer från plantor i symbios med *Rhizobium* jämfört med de utan, men att det är en signifikant skillnad i respons mellan de två sorterna. Det förhöjda innehållet av soyasapogenol korrelerar även med förekomsten av endogent jasmonat (Ranjbar Sistani *et al.* 2017).

Jasmonater är växthormon som bland annat reglerar inducerade försvarsmekanismer i respons på förekomst av herbivorer, patogener eller abiotiska stressfaktorer hos växter (Taiz *et al.* 2015). En studie på cellkulturer av lakrits (*Glycyrrhiza glabra* L.) konstaterar att förekomsten av enzymer involverade i soyasaponinsyntesen ökar och soyasaponin

ackumuleras i högre grad när exogent metyljasmonat tillförs (Hayashi *et al.* 2003). Även i ringblomma (*Calendula officinalis* L., Asteraceae) tenderar exogent metyljasmonat leda till ökad ackumulering av saponiner av oleanan-typ (Markowski *et al.* 2019).

3.3.2. Transkriptionsfaktorer

Transkriptionsfaktorer (TFs) är protein involverade i regleringen av genuttryck som svar på utvecklingsrelaterade interna signaler och externa signaler, exempelvis orsakade av stress (Yang *et al.* 2012). De kontrollerar syntes och ackumulering av sekundärmetaboliter genom att hindra eller trigga transkription av gener, ofta genom att binda till dess promotorregion, och beroende på hur den DNA-bindande regionen i proteinet ser ut kan de vara generella eller specifika sett till vilka gener de reglerar (Latchman 1997).

I tornlusern har en jasmonatinducerad transkriptionsfaktor (TF), TRITERPENE SAPONIN BIOSYNTHESIS ACTIVATING REGULATOR1 (TSAR1), inom TF-familjen basic Helix-Loop-Helix (bHLH) identifierats (Mertens *et al.* 2016). Även i kinesisk lakrits förekommer en liknande TF, bHLH3, som uppregleras av exogent jasmonat (Tamura *et al.* 2018). Överuttryck av både *TSAR1* och *bHLH3* resulterar i ökad förekomst av mRNA från enzymkodande gener med bekräftad koppling till syntesen av triterpenoida saponiner och signifikant höjda nivåer av ackumulerade soyasaponiner (Mertens *et al.* 2016; Tamura *et al.* 2018).

Tabell 4: Urval av träffar från sökning i BLAST med transkriptionsfaktorer vilka uppreglerar uttryck av gener involverade i soyasaponinsyntes.

Metod	Sökobjekt	Subjekt			
	Protein	OT %	ID %	Acc.nr.	Information
Tblast WGS	GubHLH3	88	45,8	CAADHX020001924.1	Kromosom 4
		89	44,3	CAADHX020024109.1	Kromosom 4
		100	44,5	CAADHX020009614.1	Kromosom 4
	MtTSAR1	88	66,5	CAADHX020009219.1	Kromosom 3

Acc.nr.: accessionsnummer; Gu: *Glycyrrhiza uralensis*; ID: sekvensidentitet; Mt: *Medicago truncatula*; OT: objektäckning; WGS: whole-genome shotgun.

TSAR1 och bHLH3 delar 43% aminosyraidentitet (Tamura *et al.* 2018). BLAST-resultat från sökning i WSG-sekvenser i ärta ger för båda TFs träffar på unika subjekt (Tabell 4). De tre sekvenserna kopplade till kromosom 4 (acc.nr. CAADHX020001924.1, CAADHX020024109.1 respektive CAADHX020009614.1) i ärta ger liknande objektäckning och sekvensidentitet, och är enligt information från NCBI lokaliserade nära varandra i kopplingsgruppen (acc.nr. LR535895). Sökning med enzymens översatta proteinsekvenser i NC-databaser ger träffar med objektäckning mellan 17 och 45% (Tabell B2).

4. Diskussion

4.1. Genetiska mekanismer bakom saponininnehåll i ärtä

Frågeställningen som inleder detta arbete är vad som är känt om bakomliggande genetiska mekanismer för soyasaponininnehåll i ärtä, och om hur saponinhalten varierar beroende på genotyp och miljöfaktorer.

Svaret är att trots det kommersiella intresset för ärtä och dess användning i livsmedel är syntesvägen för soyasaponiner i ärtä ännu inte klarlagd. Enbart två enzym, OSCPSY och OSCPSM, har funktionellt kopplats till soyasaponinsyntesen i ärtä (Morita *et al.* 2000; Vernoud *et al.* 2021), vilka tillhör gruppen oxidosqualene-cyklaser (OSCs). De katalyserar cykliseringen av 2,3-oxidosqualene till β -amyrin, det första steget som karaktäriserar syntesen av saponiner av oleanan-typ. Enzymet OSCPSM är multifunktionellt och katalyseringen resulterar även i andra reaktionsprodukter, främst α -amyrin (Morita *et al.* 2000).

Genen som kodar för enzymet OSCPSY, *PsBAS1*, har identifierats i en studie av Vernoud *et al.* (2021), och lokaliseras till kromosom 7 i ärtgenomet. Studien visar vidare att *PsBAS1* uttrycks i hög grad i mognande fröer, och att uttrycksmönstret korrelerar med ackumuleringen av DDMP-saponin och β -amyrin (Figur 6; 8) (Baisted 1971; Vernoud *et al.* 2021). Hos mutanter som enbart bär anlag för icke-funktionellt OSCPSY-enzym detekteras inga soyasaponiner i något av de analyserade organen (Vernoud *et al.* 2021), vilket tyder på att enzymet är avgörande för soyasaponinsyntesen i ärtä.

Övriga steg i syntesen är hydroxylering av β -amyrin till aglykonen soyasapogenol B följt av inbindning av socker och andra molekyler till aglykonen. Dessa reaktioner förmodas involvera enzym inom grupperna cytokrom P450-beroende monooxygenaser (CYPs) respektive glykosyltransferaser (GTs), vidare kan transkriptionsfaktorer (TFs) reglera uttrycket av dessa enzym. Inga proteiner ur dessa kategorier har konstaterats vara involverade i syntesen av soyasaponiner i ärtä.

Den dominerande soyasaponinen i ärtä är DDMP-saponin, även om soyasaponin Bb också påvisats (Heng *et al.* 2006a; Vernoud *et al.* 2021). Variation i saponininnehåll både i fröer av ärtä och andra baljväxter är konstaterat primärt beroende av genotyp (Shiraiwa *et al.* 1991b; Ruiz *et al.* 1995; Daveby *et al.* 1997; Heng *et al.* 2006a; Vernoud *et al.* 2021), även om abiotiska- och biotiska faktorer påverkan inte kan uteslutas.

4.2. Litteratursökning och jämförande genomik som metod

Syftet med arbetet är att undersöka vad som finns skrivet i litteraturen angående saponininnehåll i ärtor, för att underlätta vidare forskning och växtförädling. Inom familjen Fabaceae innehåller bland andra soja, tornlusern och lakrits soyasaponiner liknande de i ärtor, och syntesvägarna i dessa arter är närmare studerade (Nomura *et al.* 2007; Fukushima *et al.* 2013; Chung *et al.* 2020; Vernoud *et al.* 2021). På grund av detta har metoden inte avgränsats till att enbart beröra ärtor, utan andra arter inom Fabaceae där soyasaponiner förekommer har också studerats.

Att litteraturen rörande genetiska mekanismer kopplade till saponininnehåll i ärtor är bristfällig (Vernoud *et al.* 2021), skulle delvis kunna förklaras med att helgenomsekvensering och arts specifika genetiska verktyg för identifiering och förädling i ärtor dröjt jämfört med arter som soja och tornlusern (Schmutz *et al.* 2010; Young *et al.* 2011; Kreplak *et al.* 2019). Sett till soja kan det också bero på att produktionen globalt sett är nära tiofaldig produktionen av ärtor (Figur 1C) (FAO 2021), och att de ekonomiska incitamenten att bedriva växtförädlingsprojekt därmed är större. I lakrits är andra saponiner av oleanan-typ ekonomiskt viktiga (Fanani *et al.* 2019), vilket kan ha motiverat studier på relaterade saponiners syntesvägar.

I forskningen på saponiner i soja är, liksom för ärtor, förbättrad frökvalitet en drivkraft (Yano *et al.* 2017; Sundaramoorthy *et al.* 2019; Chung *et al.* 2020). Vad gäller forskningen på saponinsyntes i tornlusern och lakrits tycks främst andra motiv figurera. Detta kan göra att litteraturen rörande dessa arter är av lägre relevans i syfte att ge svar på frågor kring soyasaponinsyntesen i ärtfröer, trots att de evolutionärt står varandra nära (Figur 2) (Kreplak *et al.* 2019).

Flera studier konstaterar att soyasaponininnehållet varierar mellan studerade accessioner. Mätvärden för variation i soyasaponininnehåll är dock inte jämförbara sinsemellan, då olika extraherings- och analysmetoder förekommer. Vissa studier använder metoder som tenderar att degradera DDMP-saponin till soyasaponin Bb eller en proxy för soyasaponininnehållet som exempelvis soyasapogenol B (Shiraiwa *et al.* 1991a; Daveby *et al.* 1997; Bastianelli *et al.* 1998). Av denna anledning är studier med varsam extraheringsmetod mer tillförlitliga gällande fördelningen mellan de två soyasaponinerna (Heng *et al.* 2006b; Sundaramoorthy *et al.* 2019; Vernoud *et al.* 2021). Dessa studier visar att det framför allt är DDMP-saponin som ackumuleras i fröer av både soja och ärtor och att soyasaponin Bb är en nedbrytningsprodukt eller syntesföregångare. Andra felkällor som kan förekomma är testmaterialets ålder och efterskördshantering, då även lagringstid och -förhållanden visats påverka degraderingen (Daveby *et al.* 1998).

Att använda BLAST i syfte att identifiera potentiella gener med relaterad funktion är en etablerad metod (Moses *et al.* 2014b; Thimmappa *et al.* 2014; Tamura *et al.* 2018;

Vernoud *et al.* 2021). I denna studie är metoden avgränsad till att analysera sekvensernas objektäckning och identitetandel. Vilken del av de jämförda sekvenserna som är homologa har inte analyserats, vilket kan påverka nyttan av resultatet då funktion och specificitet hos de undersökta enzymgrupperna är kopplat till specifika regioner i sekvenserna (Hundle *et al.* 1992; Abe *et al.* 1993; Poralla *et al.* 1994; Kubo *et al.* 2004; Yonekura-Sakakibara & Hanada 2011; Cook *et al.* 2016). Metoden används främst för att sälla bland tillgängliga genomsekvenser i kombination med andra metoder som validerar funktionen, och det går inte att dra några slutsatser angående funktionen för homologa sekvenser i ärtan enbart utifrån sökresultat i BLAST.

Sökningar i WGS-databasen gav fler träffar och träffar av högre relevans, sett till objektäckning och sekvensidentitet, än motsvarande sökningar i NC-databaser (Tabell B2). Detta kan förklaras med att ärtas genom till stor del är utforskat (Kreplak *et al.* 2019), även gällande soyasaponinernas syntesvägar, och att potentiellt funktionella gener och mekanismer ännu inte identifierats.

Generellt ger också sökningar med den översatta proteinsekvensen (tblastn) fler träffar och träffar med högre relevans än med nukleotidsekvensen (blastn) (Tabell B2). I för funktionen avgörande regioner i proteinet kan aminosyror vara konserverade men de kodande nukleotiderna se olika ut. Detta kan förklaras med teorin om neutral molekylär evolution och att synonyma substitutioner i nukleotidsekvensen, till skillnad från icke-synonyma, inte utsätts för selektionstryck då funktionen förblir opåverkad (Choudhuri 2014). Flera studier använder proteinsekvensen från enzym eller andra protein med känd funktion i syfte att identifiera likartade sekvenser (Moses *et al.* 2014b; Tamura *et al.* 2018; Vernoud *et al.* 2021), vilket kan vara fördelaktigt om proteinsekvenser är konserverade och homologa i högre utsträckning än motsvarande nukleotidsekvenser.

Det finns inga garantier för att processer och involverade enzym är identiska eller konserverade inom den grupp där soyasaponiner förekommer. Förekomsten av liknande soyasaponiner, föregångare och intermediärer i soja och ärtan talar för att syntesen skulle kunna ta samma väg hos de båda arterna (Baisted 1971; Sundaramoorthy *et al.* 2019; Chung *et al.* 2020; Vernoud *et al.* 2021).

4.3. Tolkning av resultat från jämförande genomik

I och med den bristfälliga forskningen i ärtan inom området redogörs i resultatet för mekanismer, gener och enzym involverade i soyasaponinsyntesen i andra arter inom växtfamiljen Fabaceae. Den jämförande genomiken genomfördes i syfte att identifiera potentiella homologer till dessa sekvenser i tillgängliga protein-, nukleotid- och helgenomsekvenser i ärtan.

BLAST-resultaten för sökningar med β -amyrin-syntaser, vilka svarar för cykliseringsteget tidigt i syntesen av soyasaponiner, i arter inom ärtordningen (Fabeae) tyder på hög homologi med motsvarande identifierade enzym i ärtor (Tabell B2). Då förekomsten av OSCs i ärtas helgenomsekvens tidigare undersökts av Vernoud *et al.* (2021) med liknande metoder, exkluderas jämförelsen i detta arbete.

Cykloartenol-syntaset CASPEA (acc.nr. D89619) är homolog med de β -amyrin-syntaser som analyserats i BLAST, och även andra studier på enzym från arter inom olika fyla tyder på att OSCs är i hög grad konserverade och har liknande uppbyggnad oberoende reaktionsprodukt (Poralla *et al.* 1994; Abe & Prestwich 1995; Vishwakarma *et al.* 2013; Vernoud *et al.* 2021). Reaktioner drivna av OSCs sker tidigt i syntesvägen och kan ge upphov till en mängd primär- och sekundärmetaboliter utöver soyasaponiner (Moses *et al.* 2014a; Thimmappa *et al.* 2014). Detta kan vara en bidragande orsak till att det finns mycket litteratur angående enzymgruppens aktiva regioner och funktion, och att det i växter finns fler identifierade enzym i denna grupp jämfört med enzym involverade i senare steg i syntesvägen.

Flera enzym tillhörande gruppen CYPs har i baljväxter identifierats som involverade i hydroxyleringssteget i syntesen av soyasaponiner och andra saponiner av oleanan-typ (Tabell B1). Hydroxylas som identifierats katalysera reaktionen vid C-24 hos β -amyrin är i hög grad homologa i soja, trädgårdsböna, lakrits och tornlusern och tillhör alla underfamiljen CYP93E (Tabell B1) (Shibuya *et al.* 2006; Seki *et al.* 2008; Fukushima *et al.* 2011; Moses *et al.* 2014b).

I BLAST-resultatet bekräftas enzymet CYP93E8 (acc.nr. KF906539.1) i ärtor vara homologt med samtliga analyserade C-24-hydroxylas (Tabell 1). Sökning i WGS-databasen med CYP93E-enzymen ger även träff i en sekvens kopplad till kromosom 3 (acc.nr. CAADHX020004390.1) i ärtor. När nukleotidsekvenserna för de två subjekten jämförs i BLAST påvisas att de är 100% identiska (ej redovisade data), vilket visar att genen som kodar för enzymet CYP93E8 förmodligen lokaliseras till kromosom 3. Det krävs alternativa metoder att testa enzymets funktion som C-24-hydroxylas i ärtor, då den hittills genomförda studien påvisar låg katalytisk aktivitet vid uttryck i transgena jästceller även för andra enzym med tidigare bekräftad funktion som hydroxylas med β -amyrin som substrat (Fukushima *et al.* 2011; Moses *et al.* 2014b).

Tre enzym, CYP72A566 i kinesisk lakrits och CYP72A61v2 samt CYP72A66v2 i tornlusern, har konstaterats fungera som hydroxylas av C-22 hos β -amyrin uttryckta i jästceller (Tabell B1) (Fukushima *et al.* 2013; Tamura *et al.* 2018; Fanani *et al.* 2019). Resultat från BLAST-sökningar mot helgenomsekvenser påvisar hög objektäckning mot en nukleotidsekvens kopplad till kromosom 7 (acc.nr. CAADHX020007147.1), med en sekvensidentitet på 89 respektive 94% för de två specifika C-22-hydroxylasen (Tabell 2). Att enzymen har högre sekvensidentitet mot sekvensen i ärtor än den inbördes

sekvensidentiteten som visats vara 85% (Tamura *et al.* 2018), motiverar att det kan vara av intresse att undersöka den potentiella genen i ärtan närmare. Även för subjektsekvensen i kromosom 1 i ärtan, vilken delar viss sekvensidentitet med CYP72A-enzymen, kan vidare homologianalys vara motiverat.

BLAST-resultaten från sökningar med proteinsekvenserna för C-22-hydroxylasen i NC-databaser påvisar en identitet om cirka 50% mot CYP72A698 (acc.nr. MK534536.1) i ärtan (Tabell 2). Enligt Nelson och Werck-Reichhart (2011) har indelningen av CYPs i familjer och underfamiljer tenderat att bli otydlig och något överlappande på grund av att kategorin innefattar ett stort antal enzym. För proteinsekvensen från CYP72A69 i soja, som inte har någon dokumenterad funktion i syntesen av i ärtan påvisade saponiner, är sekvensidentiteten mot CYP72A698 i ärtan 57% (Tabell B2). Vidare är det, baserat på de specifika C-22-hydroxylasens inbördes homologi, resultaten från sökningarna i BLAST samt att ingen katalytisk aktivitet med β -amyrin som substrat påvisats (Fanani *et al.* 2019), inte möjligt att dra några slutsatser eller motivera vidare studier angående enzymets potentiella funktion i soyasaponinsyntesen i ärtan.

Båda undersökta underfamiljer av CYPs ger träffar på brassinosteroid-relaterade gener i ärtan. Brassinosteroider är steroida triterpener vilka erhålls ur cykloartenol och reglerar bland annat tillväxt och utveckling i växter (Fujioka & Yokota 2003). Växthormonen har en polycyklisk triterpenstruktur (Fujioka & Yokota 2003), och generna som förekommer som subjekt i BLAST-resultaten uttrycks i frön och groddplantor (Kang *et al.* 2001; Nomura *et al.* 2007), vilket kan förklara dess delvisa homologi med enzym involverade i syntesen av soyasaponiner.

I glykosyleringssteget i syntesen av soyasaponiner avgörs dess slutliga struktur och egenskaper. Denna del av soyasaponinsyntesen tycks vara närmare studerad i soja än i tornlusern och lakrits. I soja har GmCSyGT1, UGT73P2 och UGT91H4 identifierats som katalysatorer i inbindningen av de tre länkarna i sockerkedjan vid C-3 hos soyasaponin Bb (Figur 10). DDMP-delen vid C-22 binds in med hjälp av enzymet UGT73B4, antingen till aglykonen soyasapogenol B eller till det glykosylerade soyasaponin Bb, i bildandet av DDMP-saponin (Sundaramoorthy *et al.* 2019).

BLAST-sökningen i NC-databaser med samtliga sekvenser från analyserade GTs ger inga träffar i ärtan med en sekvensidentitet över 47% där objektäckningen är över 35% (Tabell B2). I resultatet från sökningar i WGS-databaser visas varje analyserat GT i soja ha en sekvensidentitet över 63% mot ett subjekt i ärtan vardera (Tabell 3). Dessa subjekt är dock inte unikt homologa med UGTs i soja utan har hög objektäckning och sekvensidentitet även mot övriga analyserade UGTs i tornlusern och kinesisk lakrits (Tabell B2), vilka har andra konstaterade funktioner i saponinsyntesen (Tabell B1). Enbart enzymet GmSyGT1 i soja, vilket tillhör en i analysen egen enzymkategori, har i BLAST-resultatet unika sökträffar (Tabell B2).

I sökningen i WGS-databaser med proteinsekvensen från UGT73B4 i soja och UGT73K1 i tornlusern påvisas en sekvensidentitet på 63 respektive 84% mot samma del av en sekvens på kromosom 7 (acc.nr. CAADHX020012292.1) i ärtor (Tabell 3; B2). De två enzymen delar egenskapen att kunna nyttja soyasapogenol som acceptorsubstrat och har enligt jämförelse i BLAST en inbördes proteinsekvensidentitet på 66% (ej redovisade data), men är följaktligen specifika för olika donatorsubstrat (Achnine *et al.* 2005; Sundaramoorthy *et al.* 2019). Även övriga två analyserade enzym ur familjen UGT73 visas vara homologa med den nämnda sekvensen i ärtor, om än med lägre sekvensidentitet (Tabell B2). Detta kan förklaras med att indelningen i inom UGTs är baserad på proteinernas sekvensidentitet, där gränsen för UGTs att kategoriseras inom samma familj generellt går vid 40% sekvensidentitet (Ross *et al.* 2001). Om subjektets sekvens i ärtor kodar för ett funktionellt enzym kategoriseras det därmed potentiellt inom familjen UGT73 tillsammans med andra enzym med funktion i syntesen av saponiner, varav två i syntesen av soyasaponiner i soja.

Chung *et al.* (2020) konstaterar att ett cellulosasyntas-erhållen glycosyltransferas är katalysator för inbindningen av glukuronsyra till C-3 hos aglykonen soyasapogenol B. Vidare antas denna grupp vara transmembranenzymer och kan därmed spela en avgörande roll för substratets rumsliga tillgänglighet från hydroxyleringen av membranassocierade CYPs till UGTs lokaliserade i cytosolen (Ro *et al.* 2001; Bowles *et al.* 2006; Chung *et al.* 2020; Jozwiak *et al.* 2020). Dessa resultat antyder att enzymens lokalisering på subcellulär nivå kan vara avgörande, och att andra enzymgrupper än de som huvudsakligen anses involverade i syntesen av soyasaponiner kan vara av intresse att studera.

De transkriptionsfaktorer som bekräftats uppreglera uttrycket av gener involverade i soyasaponinsyntesen, TSAR1 i tornlusern och bHLH3 i kinesisk lakrits, har båda strukturen av bHLH och en inbördes proteinsekvensidentitet på 43% (Mertens *et al.* 2016; Tamura *et al.* 2018). Den delade sekvensidentiteten som vid sökning med proteinsekvenserna i BLAST NC-databaser påvisas mot gener i ärtor, i kombination med låg objektäckning, gör resultaten irrelevanta (Tabell B2).

Funktionen av bHLH-proteinen hos kinesisk lakrits och tornlusern är motsvarande, men de gener som uppregleras av dessa TFs i respektive art har inte konstaterats ha likadana egenskaper (Mertens *et al.* 2016; Tamura *et al.* 2018). Den låga sekvensidentiteten som proteinerna delar tyder på att information om sekvensernas homologi inte nödvändigtvis förklarar funktionen för denna kategori av transkriptionsfaktorer. Sökräffar i helgenomsekvenserna kopplade till kromosom 3 och 4 i ärtor påvisar dock högre sekvensidentitet mot TSAR1 respektive bHLH3 än proteinernas inbördes identitet (Tabell 4). Detta kan motivera vidare studier på huruvida subjekten i fråga fyller någon funktion som TFs i uttrycket av soyasaponinrelaterade gener i ärtor.

4.4. Växtförädling av sorter med lägre saponininnehåll

Det finns belägg för att saponininnehållet i ärtor bidrar till fröernas och ärtproteinprodukternas oönskade smak (Heng *et al.* 2004; Heng *et al.* 2006a; Gläser *et al.* 2020). Till följd av detta efterfrågar odlare och industri sorter med lägre saponininnehåll (Hammenhag¹, personlig kommunikation), vilket motiverar att förädling på ärtor tar detta i beaktande.

I soja har konstaterats att syntesen av DDMP-saponin är kopplad till syntes av grupp-A saponin, där grupp-A saponiner ackumuleras i högre utsträckning när syntesen av DDMP-saponin hindras och vice versa (Yano *et al.* 2017; Sundaramoorthy *et al.* 2019). Detta förmodas vara på grund av att de gemensamma föregångarna tenderar att trigga uttrycket av gener involverade i omvandlingen till soyasaponiner (Sundaramoorthy *et al.* 2019), vilket gör växtförädlingsarbetet mer komplext och att det i syfte att hindra ackumuleringen av ämnen med oönskad smak är mer lovande att se till variation som reglerar syntesen i ett tidigare skede än vid hydroxyleringssteget (Yano *et al.* 2017; Sundaramoorthy *et al.* 2019). I ärtor där förekomst av grupp-A saponiner inte påvisats kan förmodas att denna mekanism inte behöver tas i beaktande, vilket talar för att förädlingsarbetet i ärtor inte har samma begränsningar som i soja.

Då det som är känt om soyasaponinsyntesen i ärtor begränsas till identifierade β -amyrinsyntaser och konstaterad förekomst av intermediärer, är det av intresse att vidare utforska vilka gener och regioner i genomet som är relevanta att jämföra. Enligt Sawai och Saito (2011) försvåras identifieringen av gener involverade i syntesen av triterpener av diversiteten bland kända enzym och enzymgrupper. Metoder som analyserar förekomsten av homologer till sekvenser av intresse, som i fallet jämförande genomik, kan vara ett sätt att filtrera bland den stora mängd data som finns tillgänglig.

Studerad litteratur visar att enbart enstaka skillnader i för funktionen avgörande regioner kan påverka om nukleotidsekvensen kodar för funktionella protein eller om aminosyrasekvensen har en funktion av intresse för syntesvägen i fråga. Med detta sagt är mer ingående homologianalys av föreslagna sekvenser i ärtor motiverat, något som inte genomförs inom ramarna för detta arbete.

Kompletterande metoder till homologianalys kan vara att studera eventuella samband i uttryck- och ackumuleringsmönster (Thimmappa *et al.* 2014). Detta genom att studera var i växten, under vilka utvecklingsstadier eller förhållanden kända gener uttrycks eller ämnen av intresse ackumuleras, och vidare undersöka om det finns gener som samuttrycks med dem. Denna metod har tidigare använts för att identifiera

¹ Cecilia Hammenhag, projektledare för projektet ”Ärtan – Garant för framtidens gröna protein” inom SLU Grogrund, i samtal med involverade aktörer under projektets årsmöte 2021-12-07.

soyasaponinrelaterade enzym i växter (Fukushima *et al.* 2011; 2013; Thimmappa *et al.* 2014; Tamura *et al.* 2018; Vernoud *et al.* 2021).

Det finns även studier som tyder på att gener som samuttrycks eller är involverade i samma syntesväg återfinns närmare varandra i genomet än vad som kan förklaras vara slumpmässig distribuering, i så kallade genkluster (Yi *et al.* 2007; Thimmappa *et al.* 2014). Detta tycks hos eukaryoter vara ett utbrett fenomen (Lee & Sonnhammer 2003) och i syntesen av triterpenoida saponiner har genkluster bland annat identifierats i havre, tornlusern, *Arabidopsis thaliana* L. och *Lotus japonicus* L. (Qi *et al.* 2004; Field *et al.* 2011; Krokida *et al.* 2013; Ribeiro *et al.* 2020). Detta är ett återkommande argument för att i fråga om identifiering av saponinrelaterade gener studera fenomenet närmare (Thimmappa *et al.* 2014; Ribeiro *et al.* 2020), att leta i för kända gener närliggande sekvenser i genomet kan därmed vara motiverat.

Utöver identifiering av potentiella gener krävs metoder att validera dess funktion. Exempel på metoder som tidigare använts i fråga om saponinsyntesen är att tysta eller överuttrycka genen i fråga i växten, där den mutanta fenotypen jämförs med vildtypen (Carelli *et al.* 2011; Mertens *et al.* 2016; Tamura *et al.* 2018; Sundaramoorthy *et al.* 2019; Confalonieri *et al.* 2021; Vernoud *et al.* 2021). Andra förekommande metoder är uttryck av genen i transgena jästceller (Fukushima *et al.* 2011; Seki *et al.* 2011; Moses *et al.* 2014b), växtcellkulturer (Hayashi *et al.* 2003; Confalonieri *et al.* 2009) eller modellväxter (Krokida *et al.* 2013). Många av dessa exempel på studier använder metoder för att studera funktionen efter eller i kombination med homologimodellering av olika slag, vilket styrker att jämförande genomik kan vara ett värdefullt verktyg.

Eftersom variation i saponininnehåll förekommer naturligt i ärtor finns det potential att använda befintligt växtmaterial att studera variationer i genomet och använda i förädlingen. Detta är fördelaktigt då sorter framtagna med tekniker som skapar variation, som transgena metoder och vissa former av inducerad mutagenes, är kontroversiella och svåra att lansera på marknaden inom Europeiska unionen, trots att viss oenighet råder (Purnhagen & Wessler 2019; Eriksson *et al.* 2020). Förädlingsprojektet på ärtor inom SLU Grogrund har tillgång till en kollektion bestående av 300 ärtaccessioner med olika genetisk bakgrund (Hammenhag, personlig kommunikation; SLU Grogrund 2021b). Detta utgör ett fördelaktigt utgångsläge i fråga om att studera variation i soyasaponininnehåll och koppla denna till genotyp, för att utveckla genetiska markörer och andra verktyg användbara inom förädlingen.

Vilka funktioner soyasaponiner har i växter är ännu inte helt utrett (Vernoud *et al.* 2021). Det finns studier som tyder på att i ärtor påvisade soyasaponiner kan vara involverade i växters inducerade försvar mot herbivorer eller patogener (Taylor *et al.* 2004; Łukasik *et al.* 2012; Ranjbar Sistani *et al.* 2017), hormonsignalering och rottillväxt (Tsurumi & Wada 1995; Tsurumi *et al.* 1996; Rahman *et al.* 2000; Rahman *et al.* 2001) och

interaktioner i rhizosfären (Waller *et al.* 1999; Ranjbar Sistani *et al.* 2017; Tsuno *et al.* 2018).

I försök att tysta uttryck av genen som kodar för C-22-hydroxylaset CYP72A61 i tornlusern erhöles inga livsdugliga embryon (Confalonieri *et al.* 2021), även andra studier påvisar reducerad tillväxt, fertilitet och grobarhet när syntesen för saponiner av oleanan-typ hindras (Naoumkina *et al.* 2010; Carelli *et al.* 2011). Detta talar för att saponiner av oleanan-typ eller saponinrelaterade gener kan spela en avgörande roll för växtens tillväxt och överlevnad. Vernoud *et al.* (2021) visar att ärtplantor utan ackumulerade soyasaponiner ger likvärdig frögroning och avkastning som vildtypen i försök i kontrollerad miljö, men konstaterar vidare att det är av intresse att studera effekterna av soyasaponiner i ärt i odling i fält.

I soja har uttrycket av DDMP-transferaset UGT73B4 visats vara i hög grad organspecifikt och kopplat till syntesen av DDMP-saponin i fröerna (Sundaramoorthy *et al.* 2019). I fröer av raps (*Brassica napus* L.) har förekomsten av glykosinolater, ämnen avgörande för insektsresistens hos korsblommiga växter men med antinutritionella egenskaper i föda för däggdjur, hindrat tillämpningen av rapsfröbaserade produkter som foder och livsmedel (So & Duncan 2021). Transportprotein som styr ackumuleringen av glykosinolater i fröer hos rapssläktingen sareptasenap (*B. juncea* (L.) Czern.) och homologa gener i raps och *A. thaliana* har identifierats (Andersen *et al.* 2013; Nambiar *et al.* 2021), vilket tros vara av framtida betydelse inom rapsförädlingen. Att identifiera organspecifika gener liknande DDMP-transferaset i soja eller transportproteinet i sareptasenap och förädla mot lägre saponininnehåll i fröer enbart, kan vara ett sätt att undvika eventuella bieffekter som låga soyasaponinhalter i samtliga organ kan föra med sig.

Att variationen i saponininnehåll är genotypberoende gör det relevant, men är även en förutsättning för, att vidare undersöka ärtas genom och genetik i syfte att ta fram sorter utifrån efterfrågan på marknaden. Att exogena faktorer påverkar saponininnehållet kan inte uteslutas (Ruiz *et al.* 1995; Vernoud *et al.* 2021), men även studerade saponinrelaterade stressresponser tycks vara sortberoende (Ranjbar Sistani *et al.* 2017). Brist på tydlig koppling mellan saponininnehåll och morfologiska egenskaper som frö- och blomfärg är även det ett argument att koppla fenotypen till genetiska markörer.

4.5. Slutsats

Med avstamp i vad som är känt angående genetiska mekanismer bakom innehållet av soyasaponiner i ärtor finns mycket kvar att undersöka innan bilden är fullständig. Tillgänglig information begränsas till det identifierade β -amyrin-syntaset och dess kodande gensekvens, uttrycksmönstret för dessa och hur ackumuleringen av soyasaponiner och dess föregångare ser ut. Att samla tillgänglig kunskap om vilka grupper av enzym som är involverade i soyasaponinsyntesen i andra växter och hur dessa enzym är uppbyggda bidrar till att öka förståelsen för hur motsvarande mekanismer i ärtor skulle kunna se ut. Att ärtas helgenomsekvens finns tillgänglig och monterad på kromosomnivå samt att variation i soyasaponininnehåll är genotypberoende skapar ytterligare förutsättningar att utforska föreslagna metoder vidare.

I syfte att uppnå målen inom Sveriges livsmedelsstrategi och de globala hållbarhetsmålen med en effektiviserad livsmedelsproduktion och en ökad andel vegetabiliska livsmedel kan baljväxter vara en del av lösningen. Lokalt är ärtor väl lämpad som alternativ proteinkälla sett till odlingstradition och ärtas odlingsegenskaper, men förekomsten av soyasaponiner är en bidragande faktor till att användningen begränsas. För att andelen ärtor som går till humankonsumtion ska öka krävs att proteinråvarans smakkvalitet förbättras och här kan växtförädling spela en avgörande roll.

Referenser

- Abe, I. & Prestwich, G.D. (1995). Identification of the active site of vertebrate oxidosqualene cyclase. *Lipids*, 30(3), 231-234. <https://doi.org/10.1007/BF02537826>
- Abe, I., Rohmer, M. & Prestwich, G.D. (1993). Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. *Chemical Reviews*, 93(6), 2189-2206. <https://doi.org/10.1021/cr00022a009>
- Aboobucker, S.I. & Suza, W.P. (2019). Why Do Plants Convert Sitosterol to Stigmasterol? *Frontiers in Plant Science*, 10(354). <https://doi.org/10.3389/fpls.2019.00354>
- Achnine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W. & Dixon, R.A. (2005). Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *The Plant Journal*, 41(6), 875-887. <https://doi.org/10.1111/j.1365-313X.2005.02344.x>
- Ali, M.M., Krishnamurthy, P., El-Hadary, M.H., Kim, J.M., Nawaz, M.A., Yang, S.H. & Chung, G. (2016). Identification and expression profiling of a new β -amyrin synthase gene (GmBAS3) from soybean. *Russian Journal of Plant Physiology*, 63(3), 383-390. <https://doi.org/10.1134/S1021443716020035>
- Andersen, T.G., Nour-Eldin, H.H., Fuller, V.L., Olsen, C.E., Burow, M. & Halkier, B.A. (2013). Integration of biosynthesis and long-distance transport establish organ-specific glucosinolate profiles in vegetative *Arabidopsis*. *The Plant Cell*, 25(8), 3133-3145. <https://doi.org/10.1105/tpc.113.110890>
- Baisted, D.J. (1971). Sterol and triterpene synthesis in the developing and germinating pea seed. *Biochemical Journal*, 124(2), 375-383. <https://doi.org/10.1042/bj1240375>
- Bastianelli, D., Grosjean, F., Peyronnet, C., Duparque, M. & Régnier, J. (1998). Feeding value of pea (*Pisum sativum*, L.) 1. Chemical composition of different categories of pea. *Animal Science*, 67, 609-619. <https://doi.org/10.1017/S1357729800033051>
- Basyuni, M., Oku, H., Inafuku, M., Baba, S., Iwasaki, H., Oshiro, K., Okabe, T., Shibuya, M. & Ebizuka, Y. (2006). Molecular cloning and functional expression of a multifunctional triterpene synthase cDNA from a mangrove species *Kandelia candel* (L.) Druce. *Phytochemistry*, 67(23), 2517-24. <https://doi.org/10.1016/j.phytochem.2006.09.016>
- Basyuni, M., Oku, H., Tsujimoto, E., Kinjo, K., Baba, S. & Takara, K. (2007). Triterpene synthases from the Okinawan mangrove tribe, Rhizophoraceae. *The FEBS Journal*, 274(19), 5028-5042. <https://doi.org/10.1111/j.1742-4658.2007.06025.x>

- Bowles, D., Lim, E.-K., Poppenberger, B. & Vaistij, F.E. (2006). Glycosyltransferases of lipophilic small molecules. *Annual Review of Plant Biology*, 57(1), 567-597. <https://doi.org/10.1146/annurev.arplant.57.032905.105429>
- Boye, J., Zare, F. & Pletch, A. (2010). Pulse proteins: Processing, characterization, functional properties and applications in food and feed. *Food Research International*, 43, 414-431. <https://doi.org/10.1016/j.foodres.2009.09.003>
- Cammareri, M., Consiglio, M.F., Pecchia, P., Corea, G., Lanzotti, V., Ibeas, J.I., Tava, A. & Conicella, C. (2008). Molecular characterization of beta-amyrin synthase from *Aster sedifolius* L. and triterpenoid saponin analysis. *Plant Science*, 175, 255-261. <https://doi.org/10.1016/j.plantsci.2008.04.004>
- Carelli, M., Biazzi, E., Panara, F., Tava, A., Scaramelli, L., Porceddu, A., Graham, N., Odoardi, M., Piano, E., Arcioni, S., May, S., Scotti, C. & Calderini, O. (2011). *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. *Plant Cell*, 23(8), 3070-81. <https://doi.org/10.1105/tpc.111.087312>
- Carrillo-Perdomo, E., Vidal, A., Kreplak, J., Duborjal, H., Leveugle, M., Duarte, J., Desmetz, C., Deulvot, C., Raffiot, B., Marget, P., Tayeh, N., Pichon, J.P., Falque, M., Martin, O.C., Burstin, J. & Aubert, G. (2020). Development of new genetic resources for faba bean (*Vicia faba* L.) breeding through the discovery of gene-based SNP markers and the construction of a high-density consensus map. *Scientific Reports*, 10(1), 6790. <https://doi.org/10.1038/s41598-020-63664-7>
- Celmeli, T., Sari, H., Canci, H., Sari, D., Adak, A., Eker, T. & Toker, C. (2018). The Nutritional Content of Common Bean (*Phaseolus vulgaris* L.) Landraces in Comparison to Modern Varieties. *Agronomy*, 8(9), 166. <https://doi.org/10.3390/agronomy8090166>
- Choudhuri, S. (2014). Chapter 2 - Fundamentals of Molecular Evolution. I: *Bioinformatics for Beginners*. Oxford: Academic Press. 27-53. <https://doi.org/10.1016/B978-0-12-410471-6.00002-5>
- Chung, S.Y., Seki, H., Fujisawa, Y., Shimoda, Y., Hiraga, S., Nomura, Y., Saito, K., Ishimoto, M. & Muranaka, T. (2020). A cellulose synthase-derived enzyme catalyses 3-O-glucuronosylation in saponin biosynthesis. *Nature Communications*, 11(1), 5664. <https://doi.org/10.1038/s41467-020-19399-0>
- Confalonieri, M., Cammareri, M., Biazzi, E., Pecchia, P., Fevèreiro, M.P.S., Balestrazzi, A., Tava, A. & Conicella, C. (2009). Enhanced triterpene saponin biosynthesis and root nodulation in transgenic barrel medic (*Medicago truncatula* Gaertn.) expressing a novel β -amyrin synthase (AsOXA1) gene. *Plant Biotechnology Journal*, 7(2), 172-182. <https://doi.org/10.1111/j.1467-7652.2008.00385.x>
- Confalonieri, M., Carelli, M., Gianoglio, S., Moglia, A., Biazzi, E. & Tava, A. (2021). CRISPR/Cas9-Mediated Targeted Mutagenesis of CYP93E2 Modulates the Triterpene Saponin Biosynthesis in *Medicago truncatula*. *Frontiers in Plant Science*, 12, 690231-690231. <https://doi.org/10.3389/fpls.2021.690231>
- Cook, D.J., Finnigan, J.D., Cook, K., Black, G.W. & Charnock, S.J. (2016). Chapter Five - Cytochromes P450: History, Classes, Catalytic Mechanism, and Industrial Application. I: Christov, C.Z. (red.) *Advances in Protein Chemistry and Structural Biology*. (105). Academic Press. 105-126. <https://doi.org/10.1016/bs.apcsb.2016.07.003>
- Corey, E.J., Matsuda, S.P. & Bartel, B. (1993). Isolation of an *Arabidopsis thaliana* gene encoding cycloartenol synthase by functional expression in a yeast mutant lacking

- lanosterol synthase by the use of a chromatographic screen. *Proceedings of the National Academy of Sciences*, 90(24), 11628. <https://doi.org/10.1073/pnas.90.24.11628>
- Daveby, Y.D., Åman, P., Betz, J.M. & Musser, S.M. (1998). Effect of storage and extraction on ratio of soyasaponin I to 2,3-dihydro-2,5-dihydroxy-6-methyl-4-pyrone-conjugated soyasaponin I in dehulled peas (*Pisum sativum* L.). *Journal of the Science of Food and Agriculture*, 78(1), 141-146. [https://doi.org/10.1002/\(SICI\)1097-0010\(199809\)78:1%3C141::AID-JSFA169%3E3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0010(199809)78:1%3C141::AID-JSFA169%3E3.0.CO;2-6)
- Daveby, Y.D., Åman, P., Betz, J.M., Musser, S.M. & Obermeyer, W.R. (1997). The Variation in Content and Changes during Development of Soyasaponin I in Dehulled Swedish Peas (*Pisum sativum* L.). *Journal of the Science of Food and Agriculture*, 73(3), 391-395. [https://doi.org/10.1002/\(SICI\)1097-0010\(199703\)73:3<391::AID-JSFA740>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0010(199703)73:3<391::AID-JSFA740>3.0.CO;2-O)
- Day, L. (2013). Proteins from land plants – Potential resources for human nutrition and food security. *Trends in Food Science & Technology*, 32(1), 25-42. <https://doi.org/10.1016/j.tifs.2013.05.005>
- De Ron, A., Sparvoli, F., Pueyo, J. & Bazile, D. (2017). Editorial: Protein Crops: Food and Feed for the Future. *Frontiers in Plant Science*, 8(105). <https://doi.org/10.3389/fpls.2017.00105>
- Doležel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysák, M.A., Nardi, L. & Obermayer, R. (1998). Plant Genome Size Estimation by Flow Cytometry: Inter-laboratory Comparison. *Annals of Botany*, 82(suppl_1), 17-26. <https://doi.org/10.1093/oxfordjournals.aob.a010312>
- Eriksson, D., Custers, R., Edvardsson Björnberg, K., Hansson, S.O., Purnhagen, K., Qaim, M., Romeis, J., Schiemann, J., Schleissing, S., Tosun, J. & Visser, R.G.F. (2020). Options to Reform the European Union Legislation on GMOs: Scope and Definitions. *Trends in Biotechnology*, 38(3), 231-234. <https://doi.org/10.1016/j.tibtech.2019.12.002>
- European Food Safety Authority (2014). *Scientific Opinion on the evaluation of allergenic foods and food ingredients for labelling purposes*. EFSA Journal. <https://www.efsa.europa.eu/en/efsajournal/pub/3894> [2021-11-11].
- Europeiska unionen (2011). *Förordning nr 1169/2011 av den 25 oktober 2011 om tillhandahållande av livsmedelsinformation till konsumenterna*. (Dokument 32011R1169) Europeiska unionens officiella tidning. <https://eur-lex.europa.eu/legal-content/SV/ALL/?uri=CELEX:32011R1169> [2021-11-11].
- Faizal, A. & Geelen, D. (2013). Saponins and their role in biological processes in plants. *Phytochemistry Reviews*, 12(4), 877-893. <https://doi.org/10.1007/s11101-013-9322-4>
- Fanani, M.Z., Fukushima, E.O., Sawai, S., Tang, J., Ishimori, M., Sudo, H., Ohyama, K., Seki, H., Saito, K. & Muranaka, T. (2019). Molecular Basis of C-30 Product Regioselectivity of Legume Oxidases Involved in High-Value Triterpenoid Biosynthesis. *Frontiers in Plant Science*, 10(1520). <https://doi.org/10.3389/fpls.2019.01520>
- Field, B., Fiston-Lavier, A.-S., Kemen, A., Geisler, K., Quesneville, H. & Osbourn, A.E. (2011). Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proceedings of the National Academy of Sciences*, 108(38), 16116-16121. <https://doi.org/10.1073/pnas.1109273108>

- Food and Agriculture Organization of the United Nations (FAO) (2021). *Production data: crops and livestock products*. <https://www.fao.org/faostat/en/#data/QCL> [2021-12-16].
- Fujioka, S. & Yokota, T. (2003). Biosynthesis and Metabolism of Brassinosteroids. *Annual Review of Plant Biology*, 54(1), 137-164. <https://doi.org/10.1146/annurev.arplant.54.031902.134921>
- Fukushima, E.O., Seki, H., Ohyama, K., Ono, E., Umemoto, N., Mizutani, M., Saito, K. & Muranaka, T. (2011). CYP716A Subfamily Members are Multifunctional Oxidases in Triterpenoid Biosynthesis. *Plant and Cell Physiology*, 52(12), 2050-2061. <https://doi.org/10.1093/pcp/pcr146>
- Fukushima, E.O., Seki, H., Sawai, S., Suzuki, M., Ohyama, K., Saito, K. & Muranaka, T. (2013). Combinatorial Biosynthesis of Legume Natural and Rare Triterpenoids in Engineered Yeast. *Plant and Cell Physiology*, 54(5), 740-749. <https://doi.org/10.1093/pcp/pct015>
- Gläser, P., Dawid, C., Meister, S., Bader-Mittermaier, S., Schott, M., Eisner, P. & Hofmann, T. (2020). Molecularization of Bitter Off-Taste Compounds in Pea-Protein Isolates (*Pisum sativum* L.). *Journal of Agricultural and Food Chemistry*, 68(38), 10374-10387. <https://doi.org/10.1021/acs.jafc.9b06663>
- Haralampidis, K., Bryan, G., Qi, X., Papadopoulou, K., Bakht, S., Melton, R. & Osbourn, A. (2001). A new class of oxidosqualene cyclases directs synthesis of antimicrobial phytoprotectants in monocots. *Proceedings of the National Academy of Sciences*, 98(23), 13431-6. <https://doi.org/10.1073/pnas.231324698>
- Hayashi, H., Huang, P. & Inoue, K. (2003). Up-regulation of Soyasaponin Biosynthesis by Methyl Jasmonate in Cultured Cells of *Glycyrrhiza glabra*. *Plant and Cell Physiology*, 44(4), 404-411. <https://doi.org/10.1093/pcp/pcg054>
- Hayashi, H., Huang, P., Kirakosyan, A., Inoue, K., Hiraoka, N., Ikeshiro, Y., Kushiro, T., Shibuya, M. & Ebizuka, Y. (2001). Cloning and characterization of a cDNA encoding beta-amyrin synthase involved in glycyrrhizin and soyasaponin biosyntheses in licorice. *Biological and Pharmaceutical Bulletin*, 24(8), 912-6. <https://doi.org/10.1248/bpb.24.912>
- Heng, L., van Koningsveld, G.A., Gruppen, H., van Boekel, M.A.J.S., Vincken, J.-P., Roozen, J.P. & Voragen, A.G.J. (2004). Protein–flavour interactions in relation to development of novel protein foods. *Trends in Food Science & Technology*, 15(3-4), 217-224. <https://doi.org/10.1016/j.tifs.2003.09.018>
- Heng, L., Vincken, J.-P., van Koningsveld, G., Legger, A., Gruppen, H., van Boekel, T., Roozen, J. & Voragen, F. (2006a). Bitterness of saponins and their content in dry peas. *Journal of the Science of Food and Agriculture*, 86(8), 1225-1231. <https://doi.org/10.1002/jsfa.2473>
- Heng, L., Vincken, J.P., Hoppe, K., van Koningsveld, G.A., Decroos, K., Gruppen, H., van Boekel, M.A.J.S. & Voragen, A.G.J. (2006b). Stability of pea DDMP saponin and the mechanism of its decomposition. *Food Chemistry*, 99(2), 326-334. <https://doi.org/10.1016/j.foodchem.2005.07.045>
- Herrera, J.B., Bartel, B., Wilson, W.K. & Matsuda, S.P. (1998). Cloning and characterization of the Arabidopsis thaliana lupeol synthase gene. *Phytochemistry*, 49(7), 1905-11. [https://doi.org/10.1016/s0031-9422\(98\)00366-5](https://doi.org/10.1016/s0031-9422(98)00366-5)
- Hostettmann, K. & Marston, A. (1995). *Saponins*. (Chemistry and pharmacology of natural products). Cambridge: Cambridge University Press.

- Huhman, D.V., Berhow, M.A. & Sumner, L.W. (2005). Quantification of Saponins in Aerial and Subterranean Tissues of *Medicago truncatula*. *Journal of Agricultural and Food Chemistry*, 53(6), 1914-1920. <https://doi.org/10.1021/jf0482663>
- Hundle, B.S., O'Brien, D.A., Alberti, M., Beyer, P. & Hearst, J.E. (1992). Functional expression of zeaxanthin glucosyltransferase from *Erwinia herbicola* and a proposed uridine diphosphate binding site. *Proceedings of the National Academy of Sciences*, 89(19), 9321-9325. <https://doi.org/10.1073/pnas.89.19.9321>
- Ito, R., Masukawa, Y. & Hoshino, T. (2013). Purification, kinetics, inhibitors and CD for recombinant β -amyrin synthase from *Euphorbia tirucalli* L and functional analysis of the DCTA motif, which is highly conserved among oxidosqualene cyclases. *The FEBS Journal*, 280(5), 1267-1280. <https://doi.org/10.1111/febs.12119>
- Iturbe-Ormaetxe, I., Haralampidis, K., Papadopoulou, K. & Osbourn, A.E. (2003). Molecular cloning and characterization of triterpene synthases from *Medicago truncatula* and *Lotus japonicus*. *Plant Molecular Biology*, 51(5), 731-743. <https://doi.org/10.1023/A:1022519709298>
- Jin, M.L., Lee, D.Y., Um, Y., Lee, J.H., Park, C.G., Jetter, R. & Kim, O.T. (2014). Isolation and characterization of an oxidosqualene cyclase gene encoding a β -amyrin synthase involved in *Polygala tenuifolia* Willd. saponin biosynthesis. *Plant Cell Reports*, 33(3), 511-519. <https://doi.org/10.1007/s00299-013-1554-7>
- Jordbruksverket (2021). *Jordbruksstatistik sammanställning 2021*. (Produktkod JO901). <https://jordbruksverket.se/om-jordbruksverket/jordbruksverkets-officiella-statistik/jordbruksverkets-statistikrapporter/statistik/2021-08-16-jordbruksstatistik---sammanstallning-2021> [2021-12-28].
- Jozwiak, A., Sonawane, P.D., Panda, S., Garagounis, C., Papadopoulou, K.K., Abebie, B., Massalha, H., Almekias-Siegl, E., Scherf, T. & Aharoni, A. (2020). Plant terpenoid metabolism co-opts a component of the cell wall biosynthesis machinery. *Nature Chemical Biology*, 16(7), 740-748. <https://doi.org/10.1038/s41589-020-0541-x>
- Kajikawa, M., Yamato, K.T., Fukuzawa, H., Sakai, Y., Uchida, H. & Ohyama, K. (2005). Cloning and characterization of a cDNA encoding beta-amyrin synthase from petroleum plant *Euphorbia tirucalli* L. *Phytochemistry*, 66(15), 1759-66. <https://doi.org/10.1016/j.phytochem.2005.05.021>
- Kang, J.-G., Yun, J., Kim, D.-H., Chung, K.-S., Fujioka, S., Kim, J.-I., Dae, H.-W., Yoshida, S., Takatsuto, S., Song, P.-S. & Park, C.-M. (2001). Light and Brassinosteroid Signals Are Integrated via a Dark-Induced Small G Protein in Etiolated Seedling Growth. *Cell*, 105(5), 625-636. [https://doi.org/10.1016/S0092-8674\(01\)00370-1](https://doi.org/10.1016/S0092-8674(01)00370-1)
- Kawano, N., Ichinose, K. & Ebizuka, Y. (2002). Molecular cloning and functional expression of cDNAs encoding oxidosqualene cyclases from *Costus speciosus*. *Biological and Pharmaceutical Bulletin*, 25(4), 477-82. <https://doi.org/10.1248/bpb.25.477>
- Kerfeld, C.A. & Scott, K.M. (2011). Using BLAST to teach "E-value-tionary" concepts. *PLoS biology*, 9(2), e1001014-e1001014. <https://doi.org/10.1371/journal.pbio.1001014>
- Kirby, J., Romanini, D.W., Paradise, E.M. & Keasling, J.D. (2008). Engineering triterpene production in *Saccharomyces cerevisiae*-beta-amyrin synthase from

- Artemisia annua*. *The FEBS Journal*, 275(8), 1852-9.
<https://doi.org/10.1111/j.1742-4658.2008.06343.x>
- Kitagawa, I., Taniyama, T., Nagahama, Y., Okubo, K., Yamauchi, F. & Yoshikawa, M. (1988). Saponin and Sapogenol. XLII. : Structures of Acetyl-soyasaponins A1, A2, and A3, Astringent Partially Acetylated Bisdesmosides of Soyasapogenol A, from American Soybean, the Seeds of *Glycine max* MERRILL. *Chemical & Pharmaceutical Bulletin* 36, 2819-2828. <https://doi.org/10.1248/CPB.36.2819>
- Kolesnikova, M.D., Wilson, W.K., Lynch, D.A., Obermeyer, A.C. & Matsuda, S.P.T. (2007). Arabidopsis Camelliol C Synthase Evolved from Enzymes That Make Pentacycles. *Organic Letters*, 9(25), 5223-5226.
<https://doi.org/10.1021/ol702399g>
- Kreplak, J., Madoui, M.-A., Cápál, P., Novák, P., Labadie, K., Aubert, G., Bayer, P.E., Gali, K.K., Syme, R.A., Main, D., Klein, A., Bérard, A., Vrbová, I., Fournier, C., d'Agata, L., Belser, C., Berrabah, W., Toegelová, H., Milec, Z., Vrána, J., Lee, H., Kougbéadjó, A., Térézol, M., Huneau, C., Turo, C.J., Mohellibi, N., Neumann, P., Falque, M., Gallardo, K., McGee, R., Tar'an, B., Bendahmane, A., Aury, J.-M., Batley, J., Le Paslier, M.-C., Ellis, N., Warkentin, T.D., Coyne, C.J., Salse, J., Edwards, D., Lichtenzweig, J., Macas, J., Doležel, J., Wincker, P. & Burstin, J. (2019). A reference genome for pea provides insight into legume genome evolution. *Nature Genetics*, 51(9), 1411-1422.
<https://doi.org/10.1038/s41588-019-0480-1>
- Krishnamurthy, P., Tsukamoto, C. & Ishimoto, M. (2020). Reconstruction of the Evolutionary Histories of UGT Gene Superfamily in Legumes Clarifies the Functional Divergence of Duplicates in Specialized Metabolism. *International Journal of Molecular Sciences*, 21(5), 1855.
<https://doi.org/10.3390/ijms21051855>
- Krokida, A., Delis, C., Geisler, K., Garagounis, C., Tsikou, D., Peña-Rodríguez, L.M., Katsarou, D., Field, B., Osbourn, A.E. & Papadopoulou, K.K. (2013). A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. *New Phytologist*, 200(3), 675-690.
<https://doi.org/10.1111/nph.12414>
- Kubo, A., Arai, Y., Nagashima, S. & Yoshikawa, T. (2004). Alteration of sugar donor specificities of plant glycosyltransferases by a single point mutation. *Archives of Biochemistry and Biophysics*, 429(2), 198-203.
<https://doi.org/10.1016/j.abb.2004.06.021>
- Kudou, S., Tonomura, M., Tsukamoto, C., Uchida, T., Sakabe, T., Tamura, N. & Okubo, K. (1993). Isolation and Structural Elucidation of DDMP-Conjugated Soyasaponins as Genuine Saponins from Soybean Seeds. *Bioscience, Biotechnology, and Biochemistry*, 57(4), 546-550.
<https://doi.org/10.1271/bbb.57.546>
- Kushiro, T., Shibuya, M. & Ebizuka, Y. (1998). Beta-amyrin synthase--cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. *European Journal of Biochemistry*, 256(1), 238-44.
<https://doi.org/10.1046/j.1432-1327.1998.2560238.x>
- Kushiro, T., Shibuya, M., Masuda, K. & Ebizuka, Y. (2000a). Mutational Studies on Triterpene Synthases: Engineering Lupeol Synthase into β -Amyrin Synthase. *Journal of the American Chemical Society*, 122(29), 6816-6824.
<https://doi.org/10.1021/ja0010709>

- Kushiro, T., Shibuya, M., Masuda, K. & Ebizuka, Y. (2000b). A novel multifunctional triterpene synthase from *Arabidopsis thaliana*. *Tetrahedron Letters*, 41(40), 7705-7710. [https://doi.org/10.1016/S0040-4039\(00\)01347-2](https://doi.org/10.1016/S0040-4039(00)01347-2)
- Latchman, D.S. (1997). Transcription factors: An overview. *The International Journal of Biochemistry & Cell Biology*, 29(12), 1305-1312. [https://doi.org/10.1016/S1357-2725\(97\)00085-X](https://doi.org/10.1016/S1357-2725(97)00085-X)
- Lee, J.M. & Sonnhammer, E.L.L. (2003). Genomic Gene Clustering Analysis of Pathways in Eukaryotes. *Genome Research*, 13(5), 875-882. <https://doi.org/10.1101/gr.737703>
- Leino, M.L. & Nygård, L. (2008). Puggor och pelusker - svenska lokalsorter av ärt. *Svensk Botanisk Tidsskrift*, 102, 153-162.
- Łukasik, I., Goławska, S., Wójcicka, A. & Sytykiewicz, H. (2012). Relationship Between Saponin Content in Alfalfa and Aphid Development. *Acta Biologica Cracoviensia s. Botanica*, vol. 54(No 2). <https://doi.org/10.2478/v10182-012-0022-y>
- Macas, J., Neumann, P. & Navrátilová, A. (2007). Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. *BMC Genomics*, 8(1), 427. <https://doi.org/10.1186/1471-2164-8-427>
- Markowski, M., Długosz, M., Szakiel, A., Durli, M., Poinson, S., Bouguet-Bonnet, S., Vernex-Loiset, L., Krier, G. & Henry, M. (2019). Increased synthesis of a new oleanane-type saponin in hairy roots of marigold (*Calendula officinalis*) after treatment with jasmonic acid. *Natural Product Research*, 33(8), 1218-1222. <https://doi.org/10.1080/14786419.2018.1460840>
- Meesapyodsuk, D., Balsevich, J., Reed, D.W. & Covello, P.S. (2007). Saponin biosynthesis in *Saponaria vaccaria*. cDNAs encoding beta-amyrin synthase and a triterpene carboxylic acid glucosyltransferase. *Plant physiology*, 143(2), 959-69. <https://doi.org/10.1104/pp.106.088484>
- Mendel, G. (1866). Versuche über Pflanzenhybriden. *Verhandlungen des naturforschenden Vereines in Brünn*, Bd. IV für das Jahr 1865, Avhandling, 3-47.
- Mertens, J., Pollier, J., Vanden Bossche, R., Lopez-Vidriero, I., Franco-Zorrilla, J.M. & Goossens, A. (2016). The bHLH Transcription Factors TSAR1 and TSAR2 Regulate Triterpene Saponin Biosynthesis in *Medicago truncatula*. *Plant physiology*, 170(1), 194-210. <https://doi.org/10.1104/pp.15.01645>
- Morita, M., Shibuya, M., Kushiro, T., Masuda, K. & Ebizuka, Y. (2000). Molecular cloning and functional expression of triterpene synthases from pea (*Pisum sativum*). *European Journal of Biochemistry*, 267(12), 3453-3460. <https://doi.org/10.1046/j.1432-1327.2000.01357.x>
- Morita, M., Shibuya, M., Lee, M.S., Sankawa, U. & Ebizuka, Y. (1997). Molecular cloning of pea cDNA encoding cycloartenol synthase and its functional expression in yeast. *Biological and Pharmaceutical Bulletin*, 20(7), 770-5. <https://doi.org/10.1248/bpb.20.770>
- Moses, T., Papadopoulou, K.K. & Osbourn, A. (2014a). Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. *Critical reviews in biochemistry and molecular biology*, 49(6), 439-462. <https://doi.org/10.3109/10409238.2014.953628>

- Moses, T., Thevelein, J.M., Goossens, A. & Pollier, J. (2014b). Comparative analysis of CYP93E proteins for improved microbial synthesis of plant triterpenoids. *Phytochemistry*, 108, 47-56. <https://doi.org/10.1016/j.phytochem.2014.10.002>
- Murray, K.E., Shipton, J., Whitfield, F.B. & Last, J.H. (1976). The volatiles of off-flavoured unblanched green peas (*Pisum sativum*). *Journal of the Science of Food and Agriculture*, 27(12), 1093-1107. <https://doi.org/10.1002/jsfa.2740271204>
- Nambiar, D.M., Kumari, J., Augustine, R., Kumar, P., Bajpai, P.K. & Bisht, N.C. (2021). GTR1 and GTR2 transporters differentially regulate tissue-specific glucosinolate contents and defence responses in the oilseed crop *Brassica juncea*. *Plant, Cell & Environment*, 44(8), 2729-2743. <https://doi.org/10.1111/pce.14072>
- Naoumkina, M.A., Modolo, L.V., Huhman, D.V., Urbanczyk-Wochniak, E., Tang, Y., Sumner, L.W. & Dixon, R.A. (2010). Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. *The Plant Cell*, 22(3), 850-866. <https://doi.org/10.1105/tpc.109.073270>
- Nationalencyklopedin (u.å.-a). *Trädgårdsärt*. <https://www.ne.se/uppslagsverk/encyklopedi/l%C3%A5ng/tr%C3%A4dg%C3%A5rds%C3%A4rt> [2021-11-08].
- Nationalencyklopedin (u.å.-b). *Regioselektivitet*. <https://www.ne.se/uppslagsverk/encyklopedi/l%C3%A5ng/regioselektivitet> [2021-12-06].
- Nelson, D. & Werck-Reichhart, D. (2011). A P450-centric view of plant evolution. *The Plant Journal*, 66(1), 194-211. <https://doi.org/10.1111/j.1365-313X.2011.04529.x>
- Nelson, D.R. (1999). Cytochrome P450 and the Individuality of Species. *Archives of Biochemistry and Biophysics*, 369(1), 1-10. <https://doi.org/10.1006/abbi.1999.1352>
- Nomura, T., Ueno, M., Yamada, Y., Takatsuto, S., Takeuchi, Y. & Yokota, T. (2007). Roles of Brassinosteroids and Related mRNAs in Pea Seed Growth and Germination. *Plant physiology*, 143(4), 1680-1688. <https://doi.org/10.1104/pp.106.093096>
- Nomura, Y., Seki, H., Suzuki, T., Ohyama, K., Mizutani, M., Kaku, T., Tamura, K., Ono, E., Horikawa, M., Sudo, H., Hayashi, H., Saito, K. & Muranaka, T. (2019). Functional specialization of UDP-glycosyltransferase 73P12 in licorice to produce a sweet triterpenoid saponin, glycyrrhizin. *The Plant Journal*, 99(6), 1127-1143. <https://doi.org/10.1111/tpj.14409>
- Oleszek, W. & Stochmal, A. (2002). Triterpene saponins and flavonoids in the seeds of *Trifolium* species. *Phytochemistry*, 61(2), 165-170. [https://doi.org/10.1016/S0031-9422\(02\)00230-3](https://doi.org/10.1016/S0031-9422(02)00230-3)
- Oleszek, W.A. (2002). Chromatographic determination of plant saponins. *Journal of Chromatography A*, 967(1), 147-162. [https://doi.org/10.1016/S0021-9673\(01\)01556-4](https://doi.org/10.1016/S0021-9673(01)01556-4)
- Osborn, A.E. (2003). Saponins in cereals. *Phytochemistry*, 62(1), 1-4. [https://doi.org/10.1016/S0031-9422\(02\)00393-X](https://doi.org/10.1016/S0031-9422(02)00393-X)
- Poralla, K., Hewelt, A., Prestwich, G.D., Abe, I., Reipen, I. & Sprenger, G. (1994). A specific amino acid repeat in squalene and oxidosqualene cyclases. *Trends in Biochemical Sciences*, 19(4), 157-158. [https://doi.org/10.1016/0968-0004\(94\)90276-3](https://doi.org/10.1016/0968-0004(94)90276-3)

- Price, K.R., Griffiths, N.M., Curl, C.L. & Fenwick, G.R. (1985). Undesirable sensory properties of the dried pea (*Pisum sativum*). The role of saponins. *Food Chemistry*, 17(2), 105-115. [https://doi.org/10.1016/0308-8146\(85\)90079-2](https://doi.org/10.1016/0308-8146(85)90079-2)
- Purnhagen, K.P. & Wesseler, J.H. (2019). Maximum vs minimum harmonization: what to expect from the institutional and legal battles in the EU on gene editing technologies. *Pest Management Science*, 75(9), 2310-2315. <https://doi.org/10.1002/ps.5367>
- Qi, X., Bakht, S., Leggett, M., Maxwell, C., Melton, R. & Osbourn, A. (2004). A gene cluster for secondary metabolism in oat: Implications for the evolution of metabolic diversity in plants. *Proceedings of the National Academy of Sciences*, 101(21), 8233-8238. <https://doi.org/10.1073/pnas.0401301101>
- Rahman, A., Ahamed, A., Amakawa, T., Goto, N. & Tsurumi, S. (2001). Chromosaponin I Specifically Interacts with AUX1 Protein in Regulating the Gravitropic Response of Arabidopsis Roots. *Plant physiology*, 125(2), 990-1000. <https://doi.org/10.1104/pp.125.2.990>
- Rahman, A., Tsurumi, S., Amakawa, T., Soga, K., Hoson, T., Goto, N. & Kamisaka, S. (2000). Involvement of Ethylene and Gibberellin Signalings in Chromosaponin I-Induced Cell Division and Cell Elongation in the Roots of Arabidopsis Seedlings. *Plant and Cell Physiology*, 41(1), 1-9. <https://doi.org/10.1093/pcp/41.1.1>
- Ranjbar Sistani, N., Kaul, H.-P., Desalegn, G. & Wienkoop, S. (2017). Rhizobium Impacts on Seed Productivity, Quality, and Protection of *Pisum sativum* upon Disease Stress Caused by *Didymella pinodes*: Phenotypic, Proteomic, and Metabolomic Traits. *Frontiers in Plant Science*, 8(1961). <https://doi.org/10.3389/fpls.2017.01961>
- Reckling, M., Bergkvist, G., Watson, C.A., Stoddard, F.L., Zander, P.M., Walker, R.L., Pristeri, A., Toncea, I. & Bachinger, J. (2016). Trade-Offs between Economic and Environmental Impacts of Introducing Legumes into Cropping Systems. *Frontiers in Plant Science*, 7(669). <https://doi.org/10.3389/fpls.2016.00669>
- Regeringen (2017). *En livsmedelsstrategi för Sverige – fler jobb och hållbar tillväxt i hela landet*. (Proposition 2016/17:104). <https://www.regeringen.se/rattsliga-dokument/proposition/2017/01/prop.-201617104/> [2021-11-09].
- Regeringen (2019). *Uppdrag att genomföra åtgärder under 2020–2025 inom ramen för livsmedelsstrategin*. (diarienummer N2019/03244/SMF). <https://www.regeringen.se/regeringsuppdrag/2019/12/uppdrag-till-sveriges-lantbruksuniversitet-att-genomfora-atgarder-under-20202025-inom-ramen-for-livsmedelsstrategin/> [2021-11-09].
- Ribeiro, B., Lacchini, E., Bicalho, K.U., Mertens, J., Arendt, P., Vanden Bossche, R., Calegario, G., Gryffroy, L., Ceulemans, E., Buitink, J., Goossens, A. & Pollier, J. (2020). A Seed-Specific Regulator of Triterpene Saponin Biosynthesis in *Medicago truncatula*. *The Plant Cell*, 32(6), 2020-2042. <https://doi.org/10.1105/tpc.19.00609>
- Rizzo, G. & Baroni, L. (2018). Soy, Soy Foods and Their Role in Vegetarian Diets. *Nutrients*, 10(1), 43. <https://doi.org/10.3390/nu10010043>
- Ro, D.K., Mah, N., Ellis, B.E. & Douglas, C.J. (2001). Functional Characterization and Subcellular Localization of Poplar (*Populus trichocarpa* × *Populus deltoides*) Cinnamate 4-Hydroxylase. *Plant physiology*, 126(1), 317-329. <https://doi.org/10.1104/pp.126.1.317>

- Roland, W.S.U., Pouvreau, L., Curran, J., van de Velde, F. & de Kok, P.M.T. (2017). Flavor Aspects of Pulse Ingredients. *Cereal Chemistry*, 94(1), 58-65. <https://doi.org/10.1094/CCHEM-06-16-0161-FI>
- Ross, J., Li, Y., Lim, E.-K. & Bowles, D.J. (2001). Higher plant glycosyltransferases. *Genome Biology*, 2(2), reviews3004.1. <https://doi.org/10.1186/gb-2001-2-2-reviews3004>
- Ruiz, R.G., Price, K.R., Rose, M.E., Arthur, A., Petterson, D.S. & Fenwick, R.G. (1995). The effect of cultivar and environment on saponin content of Australian sweet lupin seed. *Journal of the Science of Food and Agriculture*, 69, 347-351. <https://doi.org/10.1002/jsfa.2740690311>
- Sawai, S. & Saito, K. (2011). Triterpenoid Biosynthesis and Engineering in Plants. *Frontiers in Plant Science*, 2(25). <https://doi.org/10.3389/fpls.2011.00025>
- Sawai, S., Uchiyama, H., Mizuno, S., Aoki, T., Akashi, T., Ayabe, S. & Takahashi, T. (2011). Molecular characterization of an oxidosqualene cyclase that yields shionone, a unique tetracyclic triterpene ketone of *Aster tataricus*. *FEBS Letters*, 585(7), 1031-6. <https://doi.org/10.1016/j.febslet.2011.02.037>
- Sayama, T., Ono, E., Takagi, K., Takada, Y., Horikawa, M., Nakamoto, Y., Hirose, A., Sasama, H., Ohashi, M., Hasegawa, H., Terakawa, T., Kikuchi, A., Kato, S., Tatsuzaki, N., Tsukamoto, C. & Ishimoto, M. (2012). The Sg-1 glycosyltransferase locus regulates structural diversity of triterpenoid saponins of soybean. *The Plant Cell*, 24(5), 2123-2138. <https://doi.org/10.1105/tpc.111.095174>
- Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.-C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C. & Jackson, S.A. (2010). Genome sequence of the palaeopolyploid soybean. *Nature*, 463(7278), 178-183. <https://doi.org/10.1038/nature08670>
- Scholz, M., Lipinski, M., Leupold, M., Luftmann, H., Harig, L., Ofir, R., Fischer, R., Prüfer, D. & Müller, K.J. (2009). Methyl jasmonate induced accumulation of kalopanaxsaponin I in *Nigella sativa*. *Phytochemistry*, 70(4), 517-22. <https://doi.org/10.1016/j.phytochem.2009.01.018>
- Seki, H., Ohyama, K., Sawai, S., Mizutani, M., Ohnishi, T., Sudo, H., Akashi, T., Aoki, T., Saito, K. & Muranaka, T. (2008). Licorice β -amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proceedings of the National Academy of Sciences*, 105(37), 14204-14209. <https://doi.org/10.1073/pnas.0803876105>
- Seki, H., Sawai, S., Ohyama, K., Mizutani, M., Ohnishi, T., Sudo, H., Fukushima, E.O., Akashi, T., Aoki, T., Saito, K. & Muranaka, T. (2011). Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell*, 23(11), 4112-23. <https://doi.org/10.1105/tpc.110.082685>
- Semba, R.D., Ramsing, R., Rahman, N., Kraemer, K. & Bloem, M.W. (2021). Legumes as a sustainable source of protein in human diets. *Global Food Security*, 28, 100520. <https://doi.org/10.1016/j.gfs.2021.100520>

- Shepon, A., Eshel, G., Noor, E. & Milo, R. (2016). Energy and protein feed-to-food conversion efficiencies in the US and potential food security gains from dietary changes. *Environmental Research Letters*, 11(10), 105002. <https://doi.org/10.1088/1748-9326/11/10/105002>
- Shibuya, M., Hoshino, M., Katsube, Y., Hayashi, H., Kushiro, T. & Ebizuka, Y. (2006). Identification of β -amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. *The FEBS Journal*, 273(5), 948-959. <https://doi.org/10.1111/j.1742-4658.2006.05120.x>
- Shibuya, M., Katsube, Y., Otsuka, M., Zhang, H., Tansakul, P., Xiang, T. & Ebizuka, Y. (2009). Identification of a product specific beta-amyrin synthase from *Arabidopsis thaliana*. *Plant Physiology and Biochemistry*, 47(1), 26-30. <https://doi.org/10.1016/j.plaphy.2008.09.007>
- Shibuya, M., Nishimura, K., Yasuyama, N. & Ebizuka, Y. (2010). Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in *Glycine max*. *FEBS Letters*, 584(11), 2258-2264. <https://doi.org/10.1016/j.febslet.2010.03.037>
- Shiraiwa, M., Harada, K. & Okubo, K. (1991a). Composition and structure of "group B saponin" in soybean seed. *Agricultural and Biological Chemistry*, 55(4), 911-7. <https://doi.org/10.1271/bbb1961.55.911>
- Shiraiwa, M., Harada, K. & Okubo, K. (1991b). Composition and Content of Saponins in Soybean Seed According to Variety, Cultivation Year and Maturity. *Agricultural and Biological Chemistry*, 55(2), 323-331. <https://doi.org/10.1080/00021369.1991.10870575>
- Sjögren Bolin, Y. (2021). Allergi och korsallergi mot nötter, fröer, baljväxter, frukter och grönsaker. *Livsmedelsverkets rapportserie*, 05(2021). <https://www.livsmedelsverket.se/produktion-handel--kontroll/produktion-av-livsmedel/allergener/arta> [2021-11-11]
- SLU Grogrund (2021a). *SLU Grogrund - ett kunskapsnav för svensk växtförädling*. <https://www.slu.se/centrumbildningar-och-projekt/grogrund/om-grogrund/> [2021-11-06].
- SLU Grogrund (2021b). *Ärtan - garantin för framtidens gröna protein*. <https://www.slu.se/centrumbildningar-och-projekt/grogrund/projekt/artan---garantin-for-framtidens-grona-protein/> [2021-11-06].
- So, K.K.Y. & Duncan, R.W. (2021). Breeding Canola (*Brassica napus* L.) for Protein in Feed and Food. *Plants*, 10(10), 2220. <https://doi.org/10.3390/plants10102220>
- Sparg, S.G., Light, M.E. & van Staden, J. (2004). Biological activities and distribution of plant saponins. *Journal of Ethnopharmacology*, 94(2), 219-243. <https://doi.org/10.1016/j.jep.2004.05.016>
- Stacey, G. (2007). Chapter 10 - The Rhizobium-Legume Nitrogen-Fixing Symbiosis. I: Bothe, H., Ferguson, S.J. & Newton, W.E. (red.) *Biology of the Nitrogen Cycle*. Amsterdam: Elsevier. 147-163. <https://doi.org/10.1016/B978-044452857-5.50011-4>
- Sundaramoorthy, J., Park, G.T., Komagamine, K., Tsukamoto, C., Chang, J.H., Lee, J.-D., Kim, J.H., Seo, H.S. & Song, J.T. (2019). Biosynthesis of DDMP saponins in soybean is regulated by a distinct UDP-glycosyltransferase. *New Phytologist*, 222(1), 261-274. <https://doi.org/10.1111/nph.15588>
- Taiz, L., Zeiger, E., Møller, I.M. & Murphy, A.S. (2015). *Plant physiology and development*. 6:e uppl. Sunderland, Massachusetts, U.S.A: Sinauer Associates.

- Takagi, K., Nishizawa, K., Hirose, A., Kita, A. & Ishimoto, M. (2011). Manipulation of saponin biosynthesis by RNA interference-mediated silencing of β -amyrin synthase gene expression in soybean. *Plant Cell Reports*, 30(10), 1835. <https://doi.org/10.1007/s00299-011-1091-1>
- Tamura, K., Yoshida, K., Hiraoka, Y., Sakaguchi, D., Chikugo, A., Mochida, K., Kojoma, M., Mitsuda, N., Saito, K., Muranaka, T. & Seki, H. (2018). The Basic Helix–Loop–Helix Transcription Factor GubHLH3 Positively Regulates Soyasaponin Biosynthetic Genes in *Glycyrrhiza uralensis*. *Plant and Cell Physiology*, 59(4), 783-796. <https://doi.org/10.1093/pcp/pcy046>
- Taylor, W.G., Fields, P.G. & Sutherland, D.H. (2004). Insecticidal Components from Field Pea Extracts: Soyasaponins and Lysolecithins. *Journal of Agricultural and Food Chemistry*, 52(25), 7484-7490. <https://doi.org/10.1021/jf0308051>
- Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P. & Osbourn, A. (2014). Triterpene Biosynthesis in Plants. *Annual Review of Plant Biology*, 65(1), 225-257. <https://doi.org/10.1146/annurev-arplant-050312-120229>
- Touchman, J. (2010). Comparative Genomics. *Nature Education Knowledge*, 3(10). <https://www.nature.com/scitable/knowledge/library/comparative-genomics-13239404/> [2021-11-14]
- Tsuno, Y., Fujimatsu, T., Endo, K., Sugiyama, A. & Yazaki, K. (2018). Soyasaponins: A New Class of Root Exudates in Soybean (*Glycine max*). *Plant and Cell Physiology*, 59(2), 366-375. <https://doi.org/10.1093/pcp/pcx192>
- Tsurumi, S. & Wada, S. (1995). Chromosaponin I Stimulates the Elongation of Cortical Cells in Lettuce Roots. *Plant and Cell Physiology*, 36(5), 925-929. <https://doi.org/10.1093/oxfordjournals.pcp.a078840>
- Tsurumi, S., Watanabe, R. & Kamisaka, S. (1996). Chromosaponin I increases mechanical extensibility of lettuce root cell walls. *Physiologia Plantarum*, 97(4), 740-746. <https://doi.org/10.1111/j.1399-3054.1996.tb00539.x>
- Tzitzikas, E.N., Vincken, J.P., de Groot, J., Gruppen, H. & Visser, R.G. (2006). Genetic variation in pea seed globulin composition. *Journal of Agricultural and Food Chemistry*, 54(2), 425-433. <https://doi.org/10.1021/jf0519008>
- United Nations (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. (A/RES/70/1). <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication> [2021-11-09].
- Vernoud, V., Lebeigle, L., Munier, J., Marais, J., Sanchez, M., Pertuit, D., Rossin, N., Darchy, B., Aubert, G., Le Signor, C., Berdeaux, O., Lacaille-Dubois, M.-A. & Thompson, R. (2021). β -Amyrin Synthase1 Controls the Accumulation of the Major Saponins Present in Pea (*Pisum sativum*). *Plant and Cell Physiology*, 62(5), 784-797. <https://doi.org/10.1093/pcp/pcab049>
- Vishwakarma, R.K., Sonawane, P., Singh, S., Kumari, U., Ruby & Khan, B.M. (2013). Molecular characterization and differential expression studies of an oxidosqualene cyclase (OSC) gene of Brahmi (*Bacopa monniera*). *Physiology and molecular biology of plants : an international journal of functional plant biology*, 19(4), 547-553. <https://doi.org/10.1007/s12298-013-0195-1>
- Vogt, T. & Jones, P. (2000). Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends in Plant Science*, 5(9), 380-6. [https://doi.org/10.1016/s1360-1385\(00\)01720-9](https://doi.org/10.1016/s1360-1385(00)01720-9)

- Waller, G.R., Yang, C.F., Chen, L.F., Su, C.H., Liou, R.M., Wu, S.C., Young, C.C., Lee, M.R., Lee, J.S., Cheng, C.S., Chou, C.H. & Kim, D. (1999). Saponins produced during the life cycle of mungbeans and their role as allelochemicals. I: Yang, C.-R. & Tanaka, O. (red.) *Studies in Plant Science*. (6). Elsevier. 105-130. [https://doi.org/10.1016/S0928-3420\(99\)80015-9](https://doi.org/10.1016/S0928-3420(99)80015-9)
- Wang, Z., Guhling, O., Yao, R., Li, F., Yeats, T.H., Rose, J.K.C. & Jetter, R. (2011). Two oxidosqualene cyclases responsible for biosynthesis of tomato fruit cuticular triterpenoids. *Plant physiology*, 155(1), 540-552. <https://doi.org/10.1104/pp.110.162883>
- Watson, C.A., Reckling, M., Preissel, S., Bachinger, J., Bergkvist, G., Kuhlman, T., Lindström, K., Nemecek, T., Topp, C.F.E., Vanhatalo, A., Zander, P., Murphy-Bokern, D. & Stoddard, F.L. (2017). Chapter Four - Grain Legume Production and Use in European Agricultural Systems. I: Sparks, D.L. (red.) *Advances in Agronomy*. (144). Academic Press. 235-303. <https://doi.org/10.1016/bs.agron.2017.03.003>
- Yang, C.-Q., Fang, X., Wu, X.-M., Mao, Y.-B., Wang, L.-J. & Chen, X.-Y. (2012). Transcriptional Regulation of Plant Secondary Metabolism. *Journal of Integrative Plant Biology*, 54(10), 703-712. <https://doi.org/10.1111/j.1744-7909.2012.01161.x>
- Yano, R., Takagi, K., Takada, Y., Mukaiyama, K., Tsukamoto, C., Sayama, T., Kaga, A., Anai, T., Sawai, S., Ohya, K., Saito, K. & Ishimoto, M. (2017). Metabolic switching of astringent and beneficial triterpenoid saponins in soybean is achieved by a loss-of-function mutation in cytochrome P450 72A69. *The Plant Journal*, 89(3), 527-539. <https://doi.org/10.1111/tpj.13403>
- Yi, G., Sze, S.-H. & Thon, M.R. (2007). Identifying clusters of functionally related genes in genomes. *Bioinformatics*, 23(9), 1053-1060. <https://doi.org/10.1093/bioinformatics/btl673>
- Yonekura-Sakakibara, K. & Hanada, K. (2011). An evolutionary view of functional diversity in family 1 glycosyltransferases. *The Plant Journal*, 66(1), 182-193. <https://doi.org/10.1111/j.1365-313X.2011.04493.x>
- Young, N.D., Debellé, F., Oldroyd, G.E.D., Geurts, R., Cannon, S.B., Udvardi, M.K., Bénédicto, V.A., Mayer, K.F.X., Gouzy, J., Schoof, H., Van de Peer, Y., Proost, S., Cook, D.R., Meyers, B.C., Spannagl, M., Cheung, F., De Mita, S., Krishnakumar, V., Gundlach, H., Zhou, S., Mudge, J., Bharti, A.K., Murray, J.D., Naoumkina, M.A., Rosen, B., Silverstein, K.A.T., Tang, H., Rombauts, S., Zhao, P.X., Zhou, P., Barbe, V., Bardou, P., Bechner, M., Bellec, A., Berger, A., Bergès, H., Bidwell, S., Bisseling, T., Choise, N., Couloux, A., Denny, R., Deshpande, S., Dai, X., Doyle, J.J., Dubez, A.-M., Farmer, A.D., Fouteau, S., Franken, C., Gibelin, C., Gish, J., Goldstein, S., González, A.J., Green, P.J., Hallab, A., Hartog, M., Hua, A., Humphray, S.J., Jeong, D.-H., Jing, Y., Jöcker, A., Kenton, S.M., Kim, D.-J., Klee, K., Lai, H., Lang, C., Lin, S., Macmil, S.L., Magdelenat, G., Matthews, L., McCarrison, J., Monaghan, E.L., Mun, J.-H., Najar, F.Z., Nicholson, C., Noiro, C., O’Bleness, M., Paule, C.R., Poulain, J., Prion, F., Qin, B., Qu, C., Retzel, E.F., Riddle, C., Sallet, E., Samain, S., Samson, N., Sanders, I., Saurat, O., Scarpelli, C., Schiex, T., Segurens, B., Severin, A.J., Sherrier, D.J., Shi, R., Sims, S., Singer, S.R., Sinharoy, S., Sterck, L., Viollet, A., Wang, B.-B., Wang, K., Wang, M., Wang, X., Warfsmann, J., Weissenbach, J., White, D.D., White, J.D., Wiley, G.B., Wincker, P., Xing, Y., Yang, L., Yao, Z., Ying, F., Zhai,

- J., Zhou, L., Zuber, A., Dénarié, J., Dixon, R.A., May, G.D., Schwartz, D.C., Rogers, J., Quétier, F., Town, C.D. & Roe, B.A. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature*, 480(7378), 520-524. <https://doi.org/10.1038/nature10625>
- Zander, P., Amjath-Babu, T.S., Preissel, S., Reckling, M., Bues, A., Schläpke, N., Kuhlman, T., Bachinger, J., Uthes, S., Stoddard, F., Murphy-Bokern, D. & Watson, C. (2016). Grain legume decline and potential recovery in European agriculture: a review. *Agronomy for Sustainable Development*, 36(2), 26. <https://doi.org/10.1007/s13593-016-0365-y>
- Zhang, H., Shibuya, M., Yokota, S. & Ebizuka, Y. (2003). Oxidosqualene cyclases from cell suspension cultures of *Betula platyphylla* var. *japonica*: molecular evolution of oxidosqualene cyclases in higher plants. *Biological and Pharmaceutical Bulletin*, 26(5), 642-50. <https://doi.org/10.1248/bpb.26.642>

Bilagor

Tabell B1: Sammanställning av enzym som kan kopplas till syntesen av soyasaponiner eller ämnen soyasaponiner kan erhållas ur.

Typ	Enzym ^x	Art	Accessions nr.	Funktion	Referens
OSCs	OSCPY	<i>Pisum sativum</i> ^a	AB034802	β-Amyrin-syntas	(Morita <i>et al.</i> 2000; Vernoud <i>et al.</i> 2021)
	AMS1	<i>Glycine max</i> ^b	AY095999	β-Amyrin-syntas	(Takagi <i>et al.</i> 2011)
	GmBAS3	<i>Glycine max</i> ^b	KF597523	β-Amyrin-syntas	(Ali <i>et al.</i> 2016)
	GgbAS1	<i>Glycyrrhiza glabra</i> ^b	AB037203	β-Amyrin-syntas	(Hayashi <i>et al.</i> 2001)
	OSC1	<i>Lotus japonicus</i> ^b	AB181244	β-Amyrin-syntas	(Sawai <i>et al.</i> 2011)
	bAS1	<i>Medicago truncatula</i> ^b	AJ430607	β-Amyrin-syntas	(Iturbe-Ormaetxe <i>et al.</i> 2003)
	bAS	<i>Polygala tenuifolia</i> ^c	EF107623	β-Amyrin-syntas	(Jin <i>et al.</i> 2014)
	AtBAS	<i>Arabidopsis thaliana</i>	B6EXY6	β-Amyrin-syntas	(Shibuya <i>et al.</i> 2009)
	BAS	<i>Artemisia annua</i>	EU330197	β-Amyrin-syntas	(Kirby <i>et al.</i> 2008)
	OXA1	<i>Aster sedifolius</i>	AY836006	β-Amyrin-syntas	(Cammareri <i>et al.</i> 2008)
	OSCBPY	<i>Betula platyphylla</i>	AB055512	β-Amyrin-syntas	(Zhang <i>et al.</i> 2003)
	BgbAS	<i>Bruguiera gymnorhiza</i>	AB289585	β-Amyrin-syntas	(Basyuni <i>et al.</i> 2007)
	EtAS	<i>Euphorbia tirucalli</i>	AB206469	β-Amyrin-syntas	(Kajikawa <i>et al.</i> 2005)
	BAS1	<i>Nigella sativa</i>	FJ013228	β-Amyrin-syntas	(Scholz <i>et al.</i> 2009)
	OSCPNY1	<i>Panax ginseng</i>	AB009030	β-Amyrin-syntas	(Kushiro <i>et al.</i> 1998)
	OSCPNY2	<i>Panax ginseng</i>	AB014057	β-Amyrin-syntas	(Kushiro <i>et al.</i> 1998)
	ITTS1	<i>Solanum lycopersicum</i>	HQ266579	β-Amyrin-syntas	(Wang <i>et al.</i> 2011)
	BS	<i>Vaccaria hispanica</i>	DQ915167	β-Amyrin-syntas	(Meesapyodsuk <i>et al.</i> 2007)
	bAS1	<i>Avena strigosa</i> ^d	AJ311789	β-Amyrin-syntas	(Haralampidis <i>et al.</i> 2001)
	OSCPSM	<i>Pisum sativum</i> ^a	AB034803	Multifunktionellt	(Morita <i>et al.</i> 2000)
	AMY2	<i>Lotus japonicus</i> ^b	AF478455	Multifunktionellt	(Iturbe-Ormaetxe <i>et al.</i> 2003)
	CAMS1	<i>Arabidopsis thaliana</i>	At1g78955	Multifunktionellt	(Kolesnikova <i>et al.</i> 2007)
	LUP1	<i>Arabidopsis thaliana</i>	At1g78970	Multifunktionellt	(Herrera <i>et al.</i> 1998)
	LUP2	<i>Arabidopsis thaliana</i>	At1g78960	Multifunktionellt	(Kushiro <i>et al.</i> 2000b)
	KcMS	<i>Kandelia candel</i>	AB257507	Multifunktionellt	(Basyuni <i>et al.</i> 2006)
	RsM1	<i>Rhizophora stylosa</i>	AB263203	Multifunktionellt	(Basyuni <i>et al.</i> 2007)
	RsM2	<i>Rhizophora stylosa</i>	AB263204	Multifunktionellt	(Basyuni <i>et al.</i> 2007)
	TTS2	<i>Solanum lycopersicum</i>	HQ266580	Multifunktionellt	(Wang <i>et al.</i> 2011)
	CSOSC2	<i>Costus speciosus</i> ^d	AB058508	Multifunktionellt	(Kawano <i>et al.</i> 2002)
CYPs	CYP93E1	<i>Glycine max</i> ^b	AB231332	β-Amyrin-, sophoradiol C-24- hydroxylas	(Shibuya <i>et al.</i> 2006)
	CYP93E2	<i>Medicago truncatula</i> ^b	ABC59085	β-Amyrin C-24- hydroxylas	(Fukushima <i>et al.</i> 2011)
	CYP93E3	<i>Glycyrrhiza uralensis</i> ^b	AB437320	β-Amyrin C-24- hydroxylas	(Seki <i>et al.</i> 2008)

	CYP93E9	<i>Phaseolus vulgaris</i> ^b	KF906540	β-Amyrin C-24-hydroxylas	(Moses <i>et al.</i> 2014b)
	CYP72A566	<i>Glycyrrhiza uralensis</i> ^b	LC318134	β-amyirin-, 24-hydroxy-β-amyirin C-22-hydroxylas	(Tamura <i>et al.</i> 2018)
	CYP72A61v2	<i>Medicago truncatula</i> ^b	BAL45199	24-hydroxy-β-amyirin C-22-hydroxylas	(Seki <i>et al.</i> 2011; Fukushima <i>et al.</i> 2013)
	CYP72A66v2	<i>Medicago truncatula</i> ^b	MK792941	β-amyirin C-22-hydroxylas, med flera	(Fanani <i>et al.</i> 2019)
	CYP72A69	<i>Glycine max</i> ^b	LC143445	11-oxo-β-amyirin C-30-oxidas	(Yano <i>et al.</i> 2017)
GTs	UGT73B4	<i>Glycine max</i> ^b	MH255998	Soyasapogenol B, soyasaponin Bb C-22 DDMP-transferas	(Sundaramoorthy <i>et al.</i> 2019)
	UGT73K1	<i>Medicago truncatula</i> ^b	AY747626	Hederagenin-, soyasapogenol B/E β-D-Glucosyl-transferas	(Achnine <i>et al.</i> 2005)
	UGT73P2	<i>Glycine max</i> ^b	AB473730	Soyasapogenol B 3-O-glucuronid C-3 β-D-galactosyl-transferas	(Shibuya <i>et al.</i> 2010)
	UGT73P12	<i>Glycyrrhiza uralensis</i> ^b	LC315805	Glycyrrhetinsyra 3-O-monoglucuronid C-3 β-D-glucuronosyl-transferas	(Nomura <i>et al.</i> 2019)
	UGT91H4	<i>Glycine max</i> ^b	AB473731	Soyasaponin Bb' C-3 β-D-rhamnosyl-transferas	(Shibuya <i>et al.</i> 2010)
	GmCSyGT1	<i>Glycine max</i> ^b	LC500227	Soyasapogenol B/A 3-O-glucuronosyl-transferas	(Chung <i>et al.</i> 2020)
TFs	bHLH3	<i>Glycyrrhiza uralensis</i> ^b	LC318133	Uppreglering av gener kopplade till soyasaponinsyntes	(Tamura <i>et al.</i> 2018)
	TSAR1	<i>Medicago truncatula</i> ^b	KM409647	Uppreglering av gener kopplade till soyasaponinsyntes	(Mertens <i>et al.</i> 2016)

OSCs: oxidosqualene-cyklaser, CYPs: cytokrom P450-beroende monooxygenaser, GTs: glykosyltransferaser, TF: transkriptionsfaktor. ^x Enzymnamn inom gruppen OSCs är ofta förkortningar härledda ur funktionen som β-amyirin-syntas eller multifunktionellt syntas, ibland i kombination med artnamnet, UGTs och CYPs nomenklatur är baserade på proteinsekvensernas homologi, GmCSyGT1: *Glycine max* cellulosa-syntas erhållen glykosyltransferas1, bHLH3: basic Helix-Loop-Helix transcription factor 3, TSAR1: Triterpene Saponin Activating Regulator 1. ^a ärtsläktet (*Pisum*), ^b familjen ärtväxter (Fabaceae), ^c ärtordningen (Fabales), ^d monokotyledon.

Tabell B2: Sammanställning av resultat från sökning i BLAST, i databaser tillgängliga via NCBI Nucleotide Collection (NC) respektive Whole Genome Shotguns (WGS). Sökräffar presenteras med accessionsnummer och med objektäckning (%), sekvensidentitet (%) respektive E-värde inom parentes. Tomma fält indikerar att sekvenserna inte påvisats som homologa.

Sököbjekt				Subjekt (<i>Pisum sativum</i> taxid:3888)		
Enzym-grupp	Enzym	Art	Nukleotid accessionsnr.	Blastn – Nukleotid NC	Tblastn – Protein NC	Tblastn – Protein WGS ^z
OSCs	AMS1	<i>Glycine max</i> ^b	AY095999	AB034802.1 (91 / 87,0 / 0,0) AB034803.2 (92 / 84,2 / 0,0) D89619.1 (59 / 67,1 / 7e-112)	AB034802.1 (99 / 86,9 / 0,0) AB034803.2 (100 / 76,6 / 0,0) D89619.1 (98 / 56,1 / 0,0)	
	GmBAS3	<i>Glycine max</i> ^b	KF597523	AB034802.1 (98 / 73,1 / 0,0) AB034803.2 (98 / 68,9 / 0,0) D89619.1 (39 / 64,1 / 2e-29)	AB034802.1 (99 / 90,6 / 0,0) AB034803.2 (100 / 80,6 / 0,0) D89619.1 (98 / 58,5 / 0,0)	
	GgbAS1	<i>Glycyrrhiza glabra</i> ^b	AB037203	AB034802.1 (85 / 87,6 / 0,0) AB034803.2 (85 / 85,2 / 0,0) D89619.1 (66 / 66,8 / 5e-146)	AB034802.1 (99 / 90,6 / 0,0) AB034803.2 (99 / 81,5 / 0,0) D89619.1 (98 / 58,6 / 0,0)	
	OSC1	<i>Lotus japonicus</i> ^b	AB181244	AB034802.1 (99 / 85,6 / 0,0) AB034803.2 (99 / 83,6 / 0,0) D89619.1 (84 / 65,8 / 2e-136)	AB034802.1 (99 / 90,5 / 0,0) AB034803.2 (100 / 80,0 / 0,0) D89619.1 (98 / 58,5 / 0,0)	
	bAS1	<i>Medicago truncatula</i> ^b	AJ430607	AB034802.1 (92 / 90,1 / 0,0) AB034803.2 (84 / 82,6 / 0,0) D89619.1 (68 / 66,2 / 2e-139)	AB034802.1 (99 / 94,7 / 0,0) AB034803.2 (100 / 78,6 / 0,0) D89619.1 (98 / 59,0 / 0,0)	
	bAS	<i>Polygala tenuifolia</i> ^c	EF107623	AB034802.1 (77 / 77,9 / 0,0) AB034803.2 (76 / 76,1 / 0,0) D89619.1 (62 / 66,4 / 3e-130)	AB034802.1 (99 / 76,4 / 0,0) AB034803.2 (99 / 71,1 / 0,0) D89619.1 (99 / 58,4 / 0,0)	
	OXA1	<i>Aster sedifolius</i>	AY836006	AB034802.1 (98 / 73,6 / 0,0) AB034803.2 (98 / 71,2 / 0,0) D89619.1 (85 / 66,0 / 1e-127)	AB034802.1 (99 / 85,5 / 0,0) AB034803.2 (99 / 76,3 / 0,0) D89619.1 (99 / 58,1 / 0,0)	
CYPs	CYP93E1	<i>Glycine max</i> ^b	AB231332	KF906539.1 (98 / 73,9 / 0,0)	KF906539.1 (99 / 80,4 / 0,0) AF218296.1 (89 / 42,2 / 5e-81) AF532999.2	CAADHX020004390.1 (99 / 80,1 / 5e-161) CAADHX020007113.1 (99 / 42,5 / 1e-111) CAADHX020024228.1

					(74 / 44,8 / 1e-73)	(97 / 50,2 / 6e-109)
	CYP93E2	<i>Medicago truncatula</i> ^b	DQ335790	KF906539.1 (85 / 87,8 / 0,0) AF218296.1 (29 / 65,6 / 3e-27) AF175278.1 (8 / 68,5 / 9e-08) ^Y	KF906539.1 (99 / 90,4 / 0,0) AF218296.1 (89 / 43,1 / 5e-88) AF532999.2 (74 / 45,7 / 9e-73) ^Y	CAADHX020004390.1 (99 / 90,1 / 0,0) CAADHX020024228.1 (93 / 52,6 / 1e-111) CAADHX020007113.1 (99 / 42,0 / 2e-110)
	CYP93E3	<i>Glycyrrhiza uralensis</i> ^b	AB437320	KF906539.1 (99 / 81,4 / 0,0) AF532999.2 (10 / 69,1 / 1e-10) AF175278.1 (5 / 69,1 / 2e-08) ^Y	KF906539.1 (99 / 87,1 / 0,0) AF218296.1 (89 / 42,0 / 2e-90) AF532999.2 (85 / 42,1 / 7e-81) ^Y	CAADHX020004390.1 (99 / 80,2 / 4e-156) CAADHX020007113.1 (99 / 48,8 / 4e-107) CAADHX020024228.1 (90 / 51,6 / 3e-102)
	CYP93E9	<i>Phaseolus vulgaris</i> ^b	KF906540	KF906539.1 (98 / 72,6 / 0,0) AF532999.2 (10 / 69,5 1e-11) AF195812.1 (8 / 69,8 / 4e-05)	KF906539.1 (99 / 77,7 / 0,0) GU596478.1 (90 / 41,7 / 4e-85) AF218296.1 (89 / 41,8 / 4e-74) AF532999.2 (74 / 35,4 / 5e-68) ^Y	CAADHX020004390.1 (99 / 70,1 / 3e-134) CAADHX020007113.1 (99 / 37,8 / 3e-105) CAADHX020024228.1 (89 / 48,4 / 5e-97)
	CYP72A566	<i>Glycyrrhiza uralensis</i> ^b	LC318134	MK534536.1 (75 / 67,3 / 3e-83)	MK534536.1 (96 / 49,5 0,0) AB218764.1 (92 / 40,4 / 1e-135) Z49263.1 (73 / 23,7 / 2e-31) ^Y	CAADHX020024302.1 (91 / 49,1 / 4e-94) CAADHX020023846.1 (97 / 66,0 / 1e-93) CAADHX020007147.1 (97 / 88,9 / 5e-78)
	CYP72A61v2	<i>Medicago truncatula</i> ^b	AB558145.1	MK534536.1 (18 / 74,0 / 3e-42) AY883569.1 (1 / 78,7 / 2e-06) M55147.1 (2 / 74,3 / 3e-04)	MK534536.1 (92 / 51,3 1e-179) AB218764.1 (94 / 39,5 / 5e-83) Z49263.1 (28 / 36,2 / 2e-07) ^Y	CAADHX020023846.1 (96 / 63,8 / 8e-86) CAADHX020007147.1 (100 / 93,6 / 4e-79) CAADHX020024302.1 (90 / 47,2 / 1e-87)
	CYP72A66v2	<i>Medicago truncatula</i> ^b	MK792941	MK534536.1 (93 / 68,0 / 1e-131)	MK534536.1 (91 / 57,1 0,0) AB218764.1 (91 / 41,9 / 6e-132) Z49263.1 (75 / 27,5 / 8e-22) ^Y	CAADHX020007242.1 (91 / 76,0 / 1e-68) CAADHX020023846.1 (91 / 89,5 / 2e-80) CAADHX020024302.1 (81 / 57,1 / 3e-104) CAADHX020007147.1 (90 / 61,7 / 6e-53)
	CYP72A69	<i>Glycine max</i> ^b	LC143445	MK534536.1 (20 / 73,6 / 7e-58)	MK534536.1 (97 / 56,9 / 0,0) AB218764.1 (83 / 50,7 / 5e-59)	CAADHX020007242.1 (99 / 78,9 / 2e-59) CAADHX020023846.1 (95 / 64,7 / 4e-78) CAADHX020024302.1 (80 / 56,0 / 1e-99)
GTs	UGT73B4	<i>Glycine max</i> ^b	MH255998	-	KX518352.1 (66 / 45,6 / 4e-62) MF034884.1 (66 / 45,6 / 1e-61) MF034881.1 (66 / 45,9 / 2e-61) ^Y	CAADHX020012292.1 (97 / 63,1 / 0,0) CAADHX020000242.1 (95 / 47,1 / 1e-140) CAADHX020009423.1 (97 / 45,2 / 2e-137)

	UGT73K1	<i>Medicago truncatula</i> ^b	AY747626	MF034881.1 (35 / 64,4 / 4e-18) MF034884.1 (35 / 63,5 / 3e-13) KX518352.1 (25 / 64,0 / 1e-11)	MF034884.1 (99 / 36,2 / 1e-98) KX518352.1 (97 / 36,3 / 1e-98) MF034881.1 (99 / 36,9 / 5e-98)	CAADHX020012292.1 (98 / 84,2 / 0,0) CAADHX020009423.1 (99 / 47,6 / 2e-140) CAADHX020000242.1 (96 / 46,8 / 1e-140)
	UGT73P2	<i>Glycine max</i> ^b	AB473730	MF034881.1 (17 / 70,8 / 3e-26) KX518352.1 (17 / 70,8 / 3e-26) MF034884.1 (17 / 70,4 / 3e-25)	KX518352.1 (82 / 47,5 / 4e-98) MF034884.1 (82 / 47,6 / 2e-96) MF034881.1 (83 / 47,2 / 2e-90) ^Y	CAADHX020024174.1 (97 / 78,9 / 0,0) CAADHX020012692.1 (96 / 55,7 / 0,0) CAADHX020024542.1 (99 / 55,4 / 4e-172)
	UGT73P12	<i>Glycyrrhiza uralensis</i> ^b	LC315805	MF034884.1 (15 / 70,7 / 4e-24) KX518352.1 (15 / 70,7 / 4e-24) MF034881.1 (15 / 70,3 / 2e-22)	MF034881.1 (88 / 42,6 / 2e-89) MF034884.1 (85 / 45,3 / 1e-83) KX518352.1 (85 / 45,5 / 6e-83) ^Y	CAADHX020024174.1 (97 / 60,2 / 3e-163) CAADHX020012692.1 (96 / 49,7 / 1e-156) CAADHX020024542.1 (98 / 47,5 / 2e-142)
	UGT91H4	<i>Glycine max</i> ^b	AB473731	-	MF034881.1 (20 / 46,0 / 3e-18) KX518352.1 (20 / 44,4 / 3e-15) MF034884.1 (20 / 44,4 / 3e-15) ^Y	CAADHX020006075.1 (98 / 74,6 / 0,0) CAADHX020001505.1 (98 / 39,2 / 5e-109) CAADHX020008199.1 (96% / 37,1 / 1e-92)
	GmCSyGT1	<i>Glycine max</i> ^b	LC500227	-	EU681279.1 (8 / 46,6 / 1e-14)	CAADHX020000520.1 (100 / 72,6 / 3e-106) CAADHX020003300.1 (94 / 47,6 / 2e-52)
TFs	bHLH3	<i>Glycyrrhiza uralensis</i> ^b	LC318133	-	DQ399528.1 (39 / 31,8 / 3e-12) GU132941.1 (45 / 27,3 / 3e-12) GU132942.1 (45 / 27,3 / 3e-12)	CAADHX020001924.1 (88 / 45,8 / 2e-66) CAADHX020024109.1 (89 / 44,3 / 6e-66) CAADHX020009614.1 (100 / 44,5 / 2e-31)
	TSAR1	<i>Medicago truncatula</i> ^b	KM409647	-	GU132941.1 (17 / 44,8 / 1e-09) GU132942.1 (17 / 44,8 / 1e-09) DQ399528.1 (22 / 40,0 / 1e-04)	CAADHX020009219.1 (88 / 66,5 / 1e-63) CAADHX020005797.1 (59 / 36,8 / 1e-24)

^Y Sökningen gav fler träffar vilka inte presenteras i tabellen. ^Z presenterade träffar avgränsas generellt till de tre med E-värde närmast 0,0, alternativt de med en sekvensidentitet >35%. ^b familjen ärtväxter (Fabaceae), ^c ärtordningen (Fabales). OSCs: oxidosqualene-cyklaser, CYPs: cytokrom P450-beroende monooxygenaser, GTs: glykosyltransferaser, TF: transkriptionsfaktor.