Marginal lands for Growing Industrial Crops

D3.4– New industrial lines with improved quality adapted to marginal lands

Due date of deliverable: M42
Actual submission date: M52

Lead beneficiary

Wagenigen University, Netherlands

Responsible Author

Luisa Trindade WU

Additional Authors

Wolter Elbersen WR

Type			**Dissemination Level**		
R	Document, report	☒	**PU**	Public	☒
DEM	Demonstrator, pilot, prototype	☐	**CO**	Confidential, only for members of the consortium (including the Commission Services)	☐
DEC	Websites, patent fillings, videos, etc.	☐			
OTHER		☒			

Horizon 2020
European Union Funding
for Research & Innovation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No. 727698.

New industrial lines with improved quality traits

Switchgrass breeding for low lodging

Development of new materials

Switchgrass is a C4 grass from North America that has been developed also in Europe over the last 25 years. Adapted varieties have been identified for northern European conditions (UK, Netherlands, Ukraine). The variety Cave-in-Rock has generally both good persistence while delivering high yields. The main issues encountered have been lodging, especially during winter under wet and snowy conditions. On top of this it yields could be increased by increasing length of growing season which means that early re-growth in spring could also be a beneficial trait.

Selection of clones

20 plants were selected from a 20 year old switchgrass field (variety Cave-in-Rock) in Northern Groningen in the Netherlands on 27 March 2019, before spring re-growth. 15 plants were selected for low lodging upright appearance. 5 plants were selected with high lodging and leafy appearance. The variety Cave-in-Rock was selected, because it contains a high level of biodiversity and has showed a wide adaptation to very diverse - and also adverse - conditions. Lodging is a great problem as it leads to degradation of the biomass and problems in harvesting.

Figure 1. Selections were made in a 20 year old stand in Groningen, The Netherlands for low and high lodging characteristics.

The plants were grown up in the greenhouse. The 15 plants selected for low lodging upright appearance were divided in to 3 groups accor4ding to re-growth (Early, mid , late): This yielded 4 groups of clones: LL: Leafy Lodging, EU: Early Upright, MU: Mid and Upright, LU: Late and Upright.

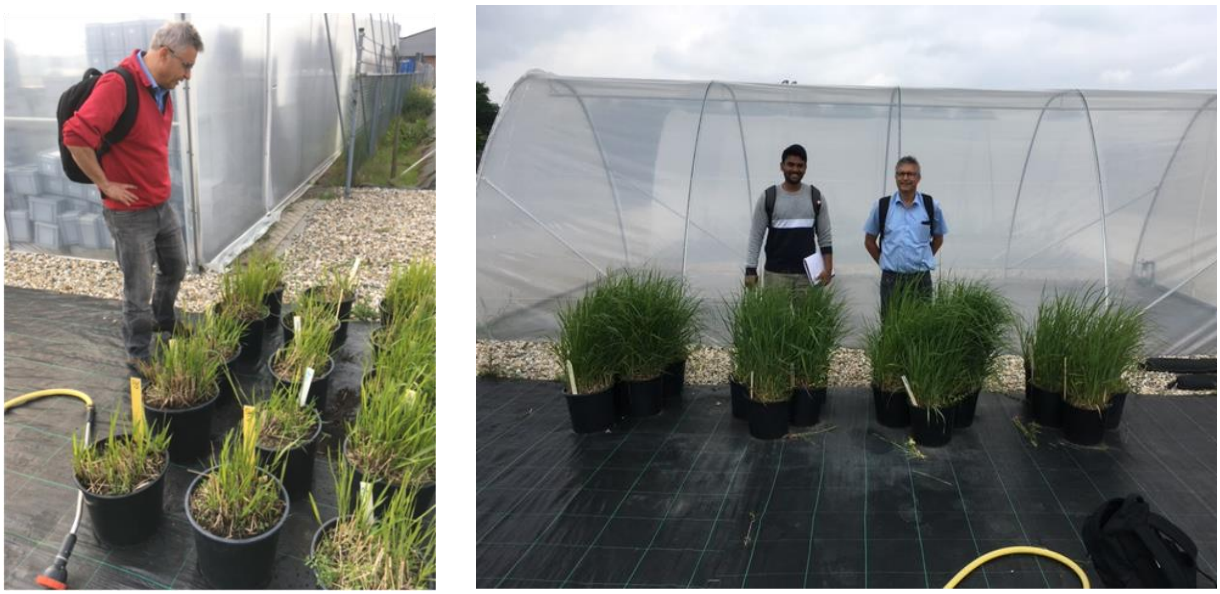

Figure 2. The selections were grouped into 4 polycross groups of 5 similar genotypes

Grouping into groups of similar clones to make new synthetic varieties

The 4 clone selections were grouped into 4 polycross groups of 5 similar genotypes according to regrowth in spring, plant height, leafiness, lodging, stem diameter and general visual similarity. These selections have been grown at Wageningen on a sandy soil with low fertility with poor water holding capacity. Within each polycross group of 5 genotypes, each genotype was divided into 4 clones and paired with each of the other genotype. This resulted in 10 pair crosses in each polycross group. In this way, a balanced composition of the offspring of the polycross groups can be made with more or less equal contributions of each of the pair crosses to the final offspring. The offspring of polycross design is a synthetic variety.

Figure 3. The 4 polycrosses (LL, EU, MU and LL) were transplanted to the field in early summer 2019.

In spring 2020 the crosses were harvested and analysed for first year DM yield and moisture content.

Evaluations of the 20 genotypes were carried out on plant height, leafiness, lodging, earliness and other traits to validate the initial selection and further characterize the genotypes. The first generation of the crosses was harvested starting in September continuing into November as seed matured.

Figure 4. Plants were monitored for maturity, height and lodging and seeds were harvested from each pair of plants starting in September by shaking of the seeds.

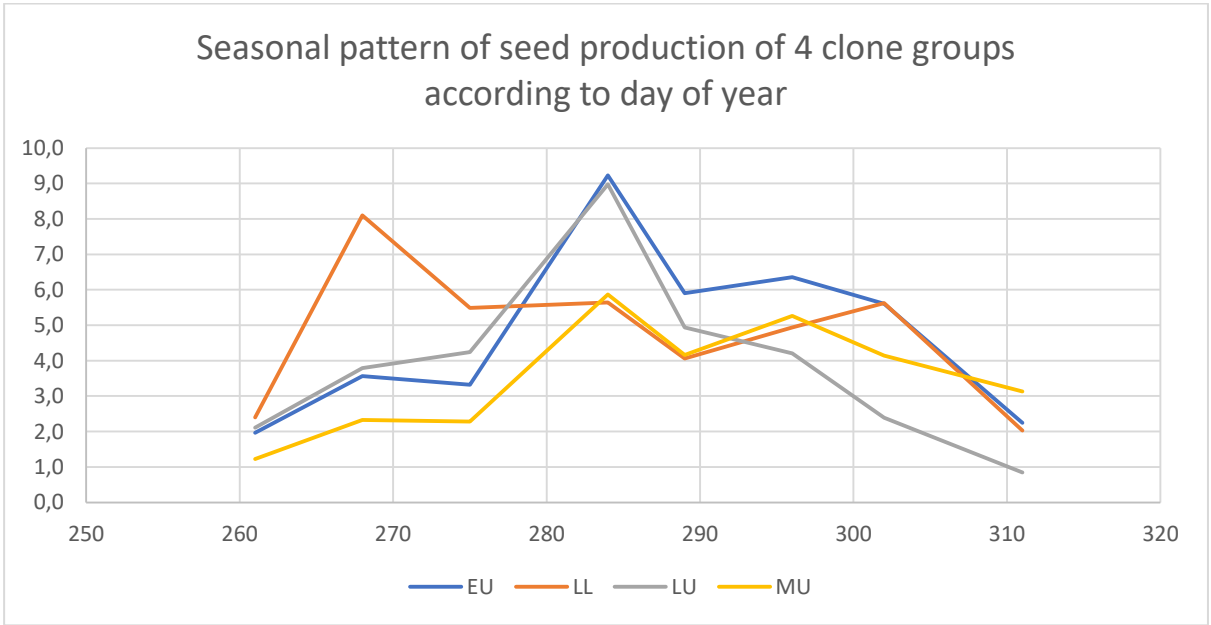

Figure 5. Seed yield over time in 2020 per selection LL: Leafy Lodging, EU: Early Upright, MU: Mid and Upright, LU: Late and Upright.

Spring dry matter yield

In order to assess the variability in early crop development, the dry matter yield above a stubble of 5 cm was determined early March 2021. The average spring dry matter yield in the first season of the four groups ranks as follows: Leafy Lodging (236 g/m^2) > Early Upright (202 g/m^2) > Mid Early Upright

(93 g/m^2 > Late Upright (133 g/m^2)(LSD=24 g/m^2). Average seed yields are positively correlated to this spring dry matter yield.

Production of seed of the SYN1 generation (plot 1 to 4.) and Phenotyping in the SYN1 generation production trial

The harvested seeds (encompassing the SYN1 populations) were stored until stored under room temperature and cold conditions and frozen for a week before transplanting into trays in the green house. The germination rate was measured (Table 1) and the seedlings were transplanted into pots until big enough for transplanting into the field at the Wageningen UR experimental farm. Per selection one clone that performed the least (low seed production, low scores in DM yield in the field appearance, low germination) was culled leaving 6 clone crosses per selection. In June 2021, 20 fields of polycrosses were established isolated from each other by more than 50 meters.

Of the four SYN1 families five replicate plots were created. Each plot consists of a equal representation of the offspring of each of the component pair crosses (6 pair crosses per SYN1, all the crosses between the four parental clones of the SYN1; this is 4 x 3 /2 crosses). Per plot (4x4 meter) some 500 plantlets were planted of 6 crosses per plot. The plots are available for making observations for phenotyping of selections and to produce the SYN2 generation. In outcrossing species like switch grass the SYN2 generation (if sufficiently uniform) can already be submitted for plant variety protection and testing for use. Table 1 lists the seed stock available after the project of the offspring of pair crosses making up the SYN1 families.

Figure 6. The seed collected from the crosses was seeded in flats for measurement of germination rate and for production of plantlets for establishment of field propagation and evaluation plots (spring 2021).

Figure 7. Establishment of SYN 2 plots for field propagation and evaluation (spring 2021).

Sequence variant analysis of the 20 clones

Leaf samples of the 20 clones were obtained. DNA was isolated and sent to BGI for whole genome shotgun sequencing (WGS, 125X coverage with paired read of 2 x 150 nucleotide length with an average insert size of about 400).

This sequence database (5.5 Tbase of total sequence length for 20 clones) has an average coverage of about 250X of the haploid genome, allowing assessment of the allele frequencies accurately also for loci that can not easily be distuinghuished on the four subgenomes in Cave-in-Rock. The reference genome sequence was based on a tetraploid variety (AR13) and Cave-in-Rock is generally considered to be an octaploid variety). That means that four pairs of homologous chromosomes occur. This means some loci might be visible with 0 to eight alleles. This makes a high coverage necessary.

The sequence database will be available for future genetic research in this switch grass material and also as a starting point for marker assisted breeding in switch grass.

To date, the first analysis of the sequence research shows a high level of heterozygosity in Cave-in-Rock clones. This is to be expected as switchgrass cannot self-pollinate (it is self-incompatible). Therefore all seed offspring is the result of a cross not a selfing and therefore the level of heterozygosity will stay high.

Figure 6 shows an example of the sequence variation as visualized in the IGV genome browser. It shows the result of read mapping of the sequence reads using the standard read mapper bwa (using 'bwa mem'). A high proportion of all reads mapped to the genome (99.6 %), but about 17 % of the read pairs did not map to the reference genome in a standard way. In some cases the inferred insert size was too large or the paired reads mapped on different contigs, indicating the reads map to positions in the reference genome that are wide apart. This is indicative of chromosomal rearrangements in our variety compared to the reference genome. As the reference genome is derived from a tetraploid, it is also possible that duplicated genomes in Cave-in-Rock have a different rearrangement than the original tetraploid genome. Despite this, the mapping result was very good with the very high proportion of reads that could be mapped.

Read mapping of all 20 clones showed a large genetic variation in Cave-in-Rock that has not been fully explored yet.

A variant calling procedure for contig JABWAI010000001.1 showed over 23,000 variant positions of Clone 1 and Clone2 compared to the reference genome. About 10,450 variants were found for which Clone 1 and Clone2 showed different genotypes (polymorphic loci in Cave-in-Rock). The number of alternative alleles per locus was hardly ever more than two on single nucleotide positions. It has to be further analysed whether using a set of single nucleotide positions large haplotype analysis can reveal whether combinations of single nucleotide position variants are part of a larger number of alleles

present. Figure 7 summarizes the numbers of homozygous and heterozygous variants in only 8 % of the genome size.

The total number of SNVs (single nucleotide variants) for which Clone 1 and Clone 2 have a different allele for the haploid whole genome size of 1.1 Gbase is estimated to be 12 million. This also includes different alleles on homologous chromosomes as reads from all these will map on the same position in the haploid genome sequence. The haploid genome sequence distinguishes between the two subgenomes in the tetraploid switchgrass, but Cave-in-Rock has a duplicated genome compared to the reference genome of AR13.

The genome sequence of this genotype AR13 and the gene annotations were used for the validation of the value of the universal QTLs for cell wall traits. Many cell wall genes of switchgrass are syntenic (have the same order on the genome) with homologous genes in other grass and cereal species (e.g. with maize) and these syntenic areas of cell wall genes prove to be linked to QTLs for cell wall traits in crops like miscanthus and maize.

Figure 6. Result of read mapping of Clone 1 and Clone 2 of Cave-in-Rock onto the reference genome of AR13.

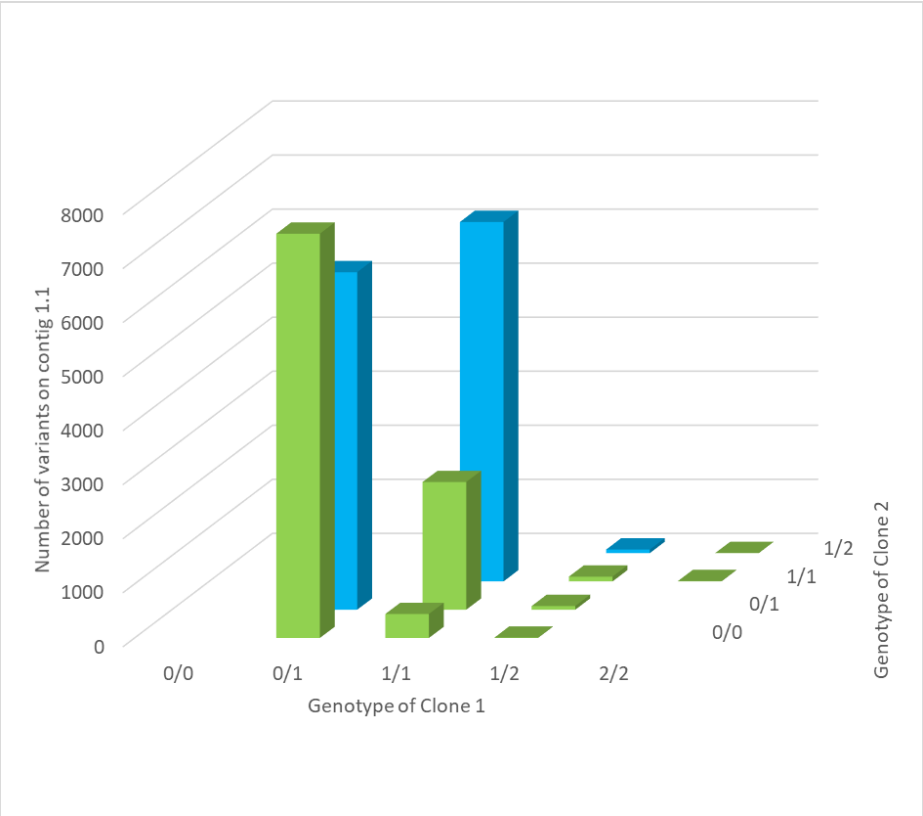

Figure 7. Sequence variants between the genetically different Clone 1 and Clone 2 selected from the switch grass accession Cave-in-Rock were analysed. The blue columns indicate were the two clones have the same genotype (differing from the reference genome sequence, either heterozygous 0/1 (so one allele equal to the reference and one allele is a variant compared to the reference) or homozygous variant (1/1) or heterozygous variant (1/2, showing two new alleles). Only very few cases occurred where two new alleles were found (when both an alternative allele 1 and 2 occurred differing from allele 0 of the reference). In45 % of the sequence positions with variants compared to the reference sequence the two clones were polymorphic

Table 1. Seed evaluation of SYN1 seed production of 5 clones per selection. Plus seeds left over after sowing and planting of the SYN1 families. (Seed material is available after the project in collaborative projects).

Selection	Clone cross		1000 seed weight (g)	Seed weight (g)	Seeds	Seed sown (g)	Approx remaining seed (g)	Precentage seedlings emerged
Leafy Lodging	22	20	2.02	43.12	21347	12.93	30.2	43%
Leafy Lodging	19	20	2.13	49.13	23066	13.63	35.5	21%
Leafy Lodging	18	19	2.18	25.1	11514	13.95	11.1	13%
Leafy Lodging	21	19	2.27	37.63	16577	14.53	23.1	29%
Leafy Lodging	18	21	2.28	28.97	12706	14.59	14.4	23%
Leafy Lodging	18	22	1.79	15.83	8844	11.46	4.4	28%
Leafy Lodging	18	20	1.94	30.36	15649	12.42	17.9	10%
Leafy Lodging	19	22	2.07	41.53	20063	13.25	28.3	40%
Leafy Lodging	21	20	2.23	52.97	23753	14.27	38.7	21%
Leafy Lodging	21	22	2.54	58.07	22862	16.26	41.8	32%
Early Upright	11	10	2.08	38.66	18587	13.31	25.3	28%
Early Upright	11	14	2.17	43.67	20124	13.89	29.8	28%
Early Upright	2	10	2.49	54.55	21908	15.94	38.6	18%
Early Upright	9	11	2.18	31.91	14638	13.95	18.0	31%
Early Upright	2	9	2.57	42.5	16537	16.45	26.1	21%
Early Upright	10	14	2.16	35.91	16625	13.82	22.1	21%
Early Upright	2	14	2.65	41.13	15521	16.96	24.2	44%
Early Upright	9	14	2.25	27.08	12036	14.40	12.7	30%
Early Upright	2	11	2.38	41.62	17487	15.23	26.4	50%
Early Upright	9	10	2.17	24.74	11401	13.89	10.9	33%
Mid Upright	16	17	2.3	21.16	9200	14.72	6.4	19%
Mid Upright	1	15	2.41	22.01	9133	15.42	6.6	37%
Mid Upright	16	4	2.11	23.54	11156	13.50	10.0	28%
Mid Upright	15	17	2.09	16.36	7828	13.38	3.0	24%
Mid Upright	16	1	2.7	46.47	17211	17.28	29.2	20%

Mid Upright	1	4	2.49	47.61	19120	15.94	31.7	28%
Mid Upright	15	4	2.14	20.61	9631	13.70	6.9	14%
Mid Upright	1	17	2.74	28	10219	17.54	10.5	23%
Mid Upright	16	15	2.06	29.64	14388	13.18	16.5	13%
Mid Upright	4	17	2.31	28.61	12385	14.78	13.8	25%
Late Upright	8	3	2.27	43.17	19018	14.53	28.6	27%
Late Upright	12	5	2.31	26.26	11368	14.78	11.5	38%
Late Upright	6	8	2.3	46.37	20161	14.72	31.7	28%
Late Upright	8	5	2.28	21.75	9539	14.59	7.2	42%
Late Upright	12	6	2.13	12.5	5869	13.63	0.0	28%
Late Upright	6	5	2.17	28.19	12991	13.89	14.3	50%
Late Upright	6	3	2.45	46.76	19086	15.68	31.1	42%
Late Upright	12	8	2.41	47.27	19614	15.42	31.8	55%
Late Upright	12	3	2.07	2.07	1000	13.25	0.0	0%
Late Upright	5	3	2.5	40.65	16260	16.00	24.7	37%

Breeding of oil/specialty crops using mutation breeding

The variety Galactica was selected as the traditional line as it was studied very well in genotype x environment interaction studies all over Europe. These studies showed that Galactica was the most robust variety with on average the highest seed yield.

The major hurdle in starting commercial production of crambe is the fact that at the seed yields attainable in short season areas and dry areas is just insufficient for economic feasibility. Crambe cannot compete in areas where winter crops (e.g. winter rapeseed) are feasible, but has a seed yield higher than in spring rapeseed, so it can compete with spring crops in short season areas. Still, for full economic feasibility either higher seed yields or a higher price for the seed or its components is needed. Two quality traits were targeted: 1) improvement of the fatty acid composition of the seed oil, 2) improvement of the feed value of the seed protein.

Crambe oil contains a high level of erucic acid (C22:1) for which a market exists in high value erucamide and other applications for very long chain fatty acids. In the separation of C22:1, a fraction of C18 fatty acids is distilled of, the so-called top fatty acids (30-40 %). With current crambe, this fraction contain about 35-50 % poly-unsaturated fatty acids, while a higher value can be obtained for this fraction if it would consist of over 80 % oleic acid.

With the aim to obtain more pure mono-unsaturated fatty acids in the oil and reduce levels of the reactive polyunsaturated fatty acids (PUFA: C18:2 + C18:3) in the oil from ~ 13% to < 10 % of total oil, crambe lines with mutations in the genes *CaFAD2-c1* and *CaFAD2-c3* were obtained.

By combining different types of mutations in Ca*FAD2-c1* and *-c3*, crambe lines with two types of altered oil composition are obtained (Table 2), respectively *FAD2* mut1, mut2 and mut3 and FAD2 mut5 and mut6. All FAD2 mutant lines flowered a week later compared to wild type cv Galactica,

produced more biomass (~ 1 tonne/ha more), showed higher seed yield (~ 500kg/ha more) and oil yield/ha (~ 100kg/ha more).
Mutant lines with oil type 1: Line FAD2 mut1, 2 and 3 are homozygous for a knockout mutation of *CaFAD2-c3* and a missense mutation in *CaFAD2-c1* (A104 to V). The seed oil of these lines are characterized by low polyunsaturated fatty acids (PUFA) (total PUFA <10%), high oleic acid (C18:1, 22%) and high erucic acid (C22:1, 60%). For breeding line FAD2 mut1 an amount of 250 kg sowing seed is available. For line FAD2 mut2 and FAD2 mut3, smaller seed batcher are available (~300-500g)
Mutant lines with oil type 2: Line FAD2 mut 4 and 5 are double knockouts of FAD2-c3 and FAD2-c1. The seed oil of this line is characterized by low polyunsaturated fatty acids (PUFA) (total PUFA <10%), high oleic acid (C18:1, 27%), a relative low level of erucic acid (C22:1, 47%) but a high level of gondoic acid (C20:1, 12%). Of line FAD2 mut5 an amount of 250 kg sowing seed is available. From line FAD2 mut4, a smaller seed batch is available (~300-500g)

Mutant lines with reduced glucosinolate levels

One of the major anti-nutritional factors in crambe seed meal is the high level of glucosinolates as this limits the use in feed ration to 5 to 10 % only. The value of the seed meal would probably double if the content of glucosinolates would drop from the current 90 µmol/g to < 18 µmol/g (dry seed weight basis).
With the aim to reduce glucosinolate levels in crambe meal from 90 µmol/g to < 18 µmol/g DW whole seed lines with knockout mutations in the target genes *CaCyp79F1-1* and *CaCyp79F1-2* were obtained. Single homozygous knockout-lines showed moderate reduced glucosinolate levels (reduction ~ 25%) but not below the target level of < 18 µmol/g DW whole seed. Only after combining knockout mutations in both genes lines with zero seed glucosinolate levels were obtained (e.g. line Cyp79F1 mut F2-02-044). Unfortunately these double knockouts (*CaCyp79F1-1*^{KO} + *CaCyp79F1-2*^{KO}) showed an aberrant “bushy” phenotype with a continuous formation of branches that do not elongate, unequal ripening, semi-sterile flowers with aberrant morphology and, most restrictive for use, seeds that don’t germinate. All together we can conclude that we targeted genes for a zero glucosinolate phenotype but, due to strong pleiotropic effects these double knockout mutants are not useful.

Mutant lines with altered oil composition + moderately reduced glucosinolate levels: To obtain breeding lines with altered oil composition and moderate reduced glucosinolate levels we crossed oil mutants (double mutants) x glucosinolate mutants (single knockout mutants). Two types of combinations were made: 1) Cyp79F1-1^{KO} x FAD2 mut5 (FAD2-c3^{KO}+FAD2-c1^{KO}) and 2) Cyp79F1-2^{KO} x FAD2 mut1 (FAD2-c3^{KO}+FAD2-c1^{A104V}). In this way each of the two oil types was combined with a single knockout in a Cyp79F1 gene for moderate glucosinolate reduction.

Mutant lines with altered oil composition plus moderate glucosinolate reduction:
For oil-type 1 (from FAD2 mut 1) a single three-double mutant was obtained.
For oil-type 2 (from FAD2 mut 5) eight three-double mutant lines were selected.

Table 2

no	line	Genotype mutant, homozygous	Remark
1	F2_01_050 F2_02_028 F2_02_037	CaCyp79F1-1^{KO}	~25% reduction in glucosinolates
2	F2_01_001 F2_01_005 F2_02_003 F2_02_006	CaCyp79F1-2^{KO}	~25% reduction in glucosinolates
3	F2_01_002 F2_01_010 F2_01_011 F2_01_037 F2-02-044 F2_01_054	CaCyp79F1-1^{KO} + CaCyp79F1-2^{KO}	Zero glucosinolate, “bushy” phenotype seeds don’t germinate
4	CA2016001 CA2016002 CA2016001 (FAD2 mutant 1, 2, 3)	FAD2-c3^{KO} + FAD2-c1^{A104V}	Mutant Oil-type 1
5	CA2016004 CA2016005 (FAD2 mutant 4, 5)	Fad2-c3^{KO} + FAD2c1^{KO}	Mutant Oil-type 2
6	F3_14_006_002 F3_14_006_006 F3_14_006_015 F3_14_006_019 F3_14_006_003 F3_14_040_009 F3_14_040_011 F3_14_040_018	Cyp791-2^{KO} + Fad2-c3^{KO}+FAD2-c1^{A104V}	Mutant Oil type 1 and moderate glucosinolate reduction
7	F2_07_18	Cyp79F1-1^{KO} + Fad2-c3^{KO}+ FAD2-c1^{KO}	Mutant Oil-type 2 and moderate glucosinolate reduction

Table 3: Delivery 3.5 mutant lines with altered oil composition and/or glucosinolate reduction

Fatty acid composition, % of total oil	**WT crambe**	**Oil-type 1 FAD2 Mutant 1**	**Oil-type 2 FAD2 Mutant 5**
PUFA (C18:2+C18:3)	14%	8% ↓- 42%	7% ↓- 50%
Oleic acid (C18:1)	16%	22% ↑+ 38%	27% ↑+ 69%
Eicosadienoic (C20:1)	3%	3%	12%↑+ 300%
Erucic acid (C22:1)	56%	60%↑+ 7%	47% ↓+ 16%
Nervonic acid (C24:1)	1%	1%	1%
others	9%	6%	5%

Table 2: Fatty acid composition of two mutant oil-types of *Crambe abyssinica* cv Galactica