Online Resource 4 - Examples of lab report fragments for each rubric score per item (translated from Dutch)

INTRODUCTION	
Rubric Score	**Lab Report fragment**
Insufficient *The research goal of the experiment is not provided.*	Multiple techniques were used to determine whether a gene is located on the chromosome
Sufficient *The research goal of the experiment is not described in own words and/or an explanation on gene mapping is absent.*	In this experiment we will determine the location of a mutation on the chromosome of Arabidopsis. For this, we use gene mapping. For gene mapping we run a PCR with different primers for InDel marker genes. These genes differ in size between two ecotypes, Le ren Col-0, between which we did a crossing that resulted in the plants we used for gene mapping. Col-0 and Ler are the P1 crossing, Ler carries the homozygous mutation. The crossing results in the F1 generation. We use the F2 generation that originated through self-pollination of heterozygous F1 plants. The PCR results in a pattern of bands for which we can determine whether recombination occurred. The results of the PCR will be compared within the whole group [of students], and with that data we can determine the distance of the mutation to the markers.
Excellent *The introduction is described in own words and contains an explanation on gene mapping.*	Mutations can result in different phenotypes of *Arabidopsis thaliana*. Mutations can result in different angles between stem and branches, or excessive flowering on branches that already have a flower. The position of the mutation on the genome can be determined with *linkage mapping*. This is a technique wherein markers with known genomic positions are used to determine the location of the mutated gene. Henceforth, different accessions, such as Landsberg (Ler) and Columbia (Col-0), are used to determine how often recombination takes place between the mutation and the marker. A homozygous mutant Ler and a homozygous wildtype Col-0 form the P generation are crossed. This results in a heterozygous F1 generation, which reproduces through self-pollination. The resulting F2-population consists of homozygous wildtype, heterozygous and homozygous mutants. The chromosomes are however a mix of Col-0 and Ler varieties because of recombination during meiosis. Consequently, an indication of the genomic position of the can be given. When the marker and the mutation are near each other, then it is expected that recombination between the marker and the mutation does not often occur because the chance of this to happen is small. If the mutation and marker are further apart, this chance will be higher. To determine this, F2 plants with a homozygous mutation should be selected based on their phenotype. Henceforth, one can determine whether recombination occurred between the marker and the mutation. The original mutation was positioned on the Ler accession. If a plant with a homozygous mutation and the marker of Col-0 is found, then recombination occured between the marker and the mutation. This can be determined with a PCR and subsequent gel electrophoresis. Whether this event occurred can be determined with many F2 plants, whereby a genotype frequency can be determined. First it needs to be determined

	how many mutants have a Ler genotype, and how many a Col-0 genotype. The frequency, and thus an indication of the genomic position, can be determined by dividing the number Col-0 fragments with the total number of fragments. The marker with the lowest recombination frequency, is closest to the mutation because they are more often inherited together. In this experiment, a mutant is used with increased flower growth, also on branches that already have flowers.

METHODS	
Rubric Score	**Lab Report fragment**
Insufficient *The Methods section is absent.*	According to experiment 13 on page 29 and 30 of the manual.
Sufficient *The method section refers to the lab protocol but does not contain any adjustments from the protocol and/or descriptions of the specific content of the mastermix.*	The manual is retained, and steps are pursued (page 29 and 30). 5 mutant F2 plants were selected instead of 8 mutants. Several steps were pursued differently on day 1 "Quick DNA extraction for PCR". For step 2 (see manual), the entirety was incubated for 45 minutes. For step 3, centrifugation occurred for 10 minutes. For step 4, centrifugation occurred for 10 minutes. For step 5, centrifugation occurred for 10 minutes. For step 6, after adding ethanol, another centrifugation occurred for 10 minutes. For step 7, MQ water was used. Materials - Arabidopsis thaliana F2 population - Arabidopsis thaliana Col and Ler F1 plants - DNA isolation materials - PCR en DNA electrophoresis materials

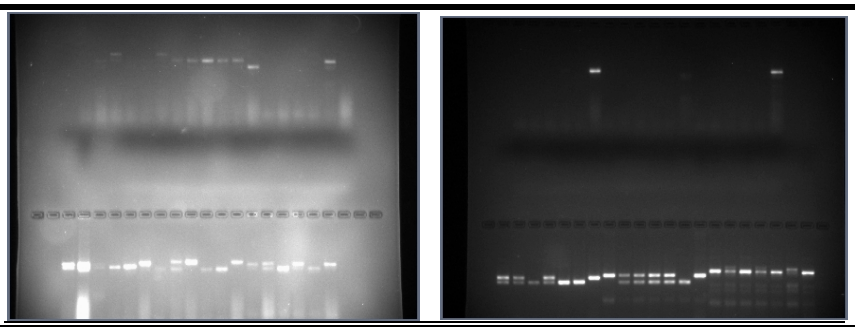
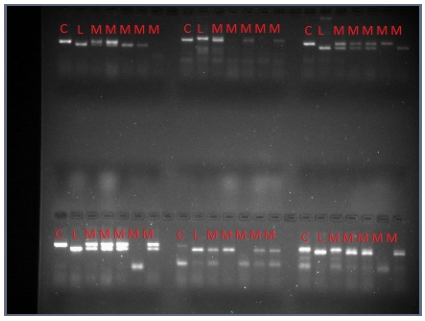
Excellent
Adjustments from the protocol and the specific content of the mastermix are provided.
	The experiment was performed according to the protocol and with the materials described on pages 29 and 30 of the lab manual. Appendix D and E V1-4 were also used (p48-52).
	• 5 mutant F2 plants were used instead of 8.
	• Adjustments Appendix E V1:
	  ○ In step 1, only 1/3 leaf was used per Eppendorf tube.
	  ○ In step 2, only 45 minutes of incubation was performed.
	  ○ In step 3, only 10 min of centrifugation was performed.
	  ○ In step 4, mixing means vortexing. Incubation took place for 10 min at room temperature.
	  ○ In step 5, only 10 min of centrifugation occurred.
	  ○ In step 6, 250 µl EtOH was added and centrifuged for 10 min. Next, the EtOH was removed with the pipet and again centrifuged for 5 sec to remove the final EtOH.
	  ○ In step 7, 650 µl miliQ was added instead of TE.
	• 9 mastermixes were made for 9 primers. Every mastermix contained a multiplication of 8 for 7 PCRs.
	• The quantities of the mastermix are somewhat different than described in Appendix E V3 on pages 49 and 50 of the lab manual. Ever mastermix contains 8 x 45 µl:

- 40 µl Taq buffer (x10).
- 4 µl dNTP's (10mM).
- 16 µl primers (forward and reverse together).
- 8 µl Taq (diluted).
- 292 µl water.

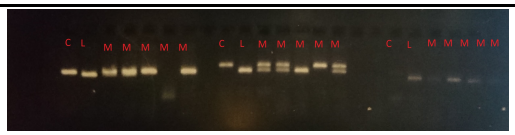
- Finally, the Taq buffer was added to prevent formation of dimers.
- 5 µl of DNA was pipetted per well, with new DNA in every row. On the day of the gel electrophoresis, 10 µl loading colour buffer was added in every well on the plate.
- During gel electrophoresis, 25 µl of PCR mix was added to the gel. An 8 Channel Micropipette was used for pipetting.
- The running of the gel took place at 125 Volt for 45 minutes.

We needed to fill 9 x 7 = 63 wells in total with the PCR mix. The DNA of the plants was pipetted in rows (so in total 7 rows). The mastermixes with the different primers were pipetted per column (so 9 columns). However, this went wrong for the first mastermix, cause the same mastermix was pipetted over the entire first row. A new mastermix was then made and new DNA of *Col* was used. The first row was not used anymore, and *Col* is the last row now.

RESULTS

Rubric Score	**Lab Report fragment**
Insufficient *The results are absent or only contain pictures of the gel electrophoresis.*	
Sufficient *One or more of the following results are absent:* – *Phenotypic description on the mutants* – *Genotypic description of own samples* – *Genotypic description on*	Wildtype plants generally have a longer stem than mutant plants, and their leaves are oriented with a 90-degree angle from the stem. The leaves of the mutants point downwards, which make them recognizable. 

samples from fellow students



Scores:
C - gene of Col-0 wildtype
L - gene of Ler mutant
H - heterozygous
X – cannot score

Table 12:

	Mutant 1	**Mutant 2**	**Mutant 3**	**Mutant 4**	**Mutant 5**
ind4-12	**C**	**C**	**L**	**L**	**X**
ciw-4	**C**	**X**	**C**	**X**	**C**
ind2-40	**H**	**H**	**H**	**C**	**L**
k18p6	**H**	**H**	**H**	**X**	**H**
ind4-6	**H**	**C**	**L**	**H**	**H**
ciw9	**C**	**L**	**L**	**X**	**L**
t1b9	**C**	**C**	**C**	**X**	**C**
f21m12	**H**	**H**	**L**	**C**	**H**
nga111	**L**	**L**	**X**	**X**	**X**

Excellent
The results contain the following:
- *Phenotypic description on the mutants*
- *Genotypic description of own samples*
- *Genotypic description on the samples from fellow students*

The mutant in this experiment is landsberg erecta, which can be recognized by its disturbed inflorescence. The mutation results in a recessive phenotype. Thus you need to use the mutant F2 plants for mapping to be sure that you've got the mutation.

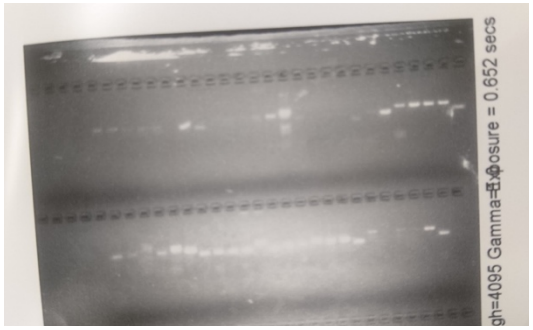


Figure 13 the results of the gel electrophorese of F2 and parents. The 9 bands on the upper right are our own data.

Loading the gel started on the upper right. Nine different primers were tested, and every primer was tested for 6 different plant DNA; thus 6 wells per primer. 1 parent wildtype Col, 1 parent mutant Ler and 4 mutant F2. First primer 1 with plant DNA was loaded, then primer 2, and

so on. As such, for every allele it can be determined if the mutant contains the wildtype allele, the mutant allele, or both. Bands were not visible in all wells. For example, primer 1 K18P6 is F2 3 times Ler and 1 times Col.

	P1 [F21M12]	P1 [nga 111]	P1 [Ind2-40]	P1 [T1B9]	P1 [Ciw 4]	P1 [Ind4-6]	P1 [Ind4-12]	P1 [K18P6]	P1 [Ciw9]
C	13	10	20	33	20	10	18	19	
L	19	15	25	13	5	7	15	17	
H	16	6	23	19	14	1	28	34	

Table 2 Results of the whole group, showing how often homozygous C or L was detected.

The mapping results of the whole laboratory group are presented in table 2. The numbers show how often a band was detected in mutants for Col, Ler or both. In this final case it is marked with a H (heterozygous).

	CHR 1	CHR 1	CHR 2	CHR 3	CHR 3	CHR 4	CHR 4	CHR 5	CHR 5
	F21 M12	Nga 111	Ind2 -40	T1B 9	Ciw 4	Ind4 -6	Ind4 -12	K18 P6	Ciw 9
C	42	21	63	85	54	21	64	72	32
L	54	36	75	45	24	15	58	68	66
Recom %	43,8	36,9	44,7	65,4	69,2	58,3	52,5	51,4	32,7

Table 3 recombination percentages calculated from the whole group

Recombination percentages are calculated in table 3. For this, the number of alleles were first calculated for Col and Ler. C or L indicates homozygous, thus two alleles. H indicates heterozygous, thus 1 allele L and 1 C. Results from table 2 are used for every calculation. For example, the number of C alleles for F21M12= 13*2+16= 42. The outcomes are then used to calculate recombination frequencies. This was done by dividing the number of C alleles with the total number of alleles. For example, for F21M12 this was (42/42+54)*100=43,8%.

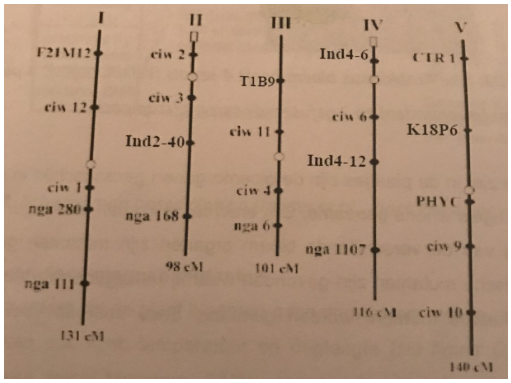
The distance in Cm was then calculated with the following formula:
D = 25 x ln[(100+2r)/(100–2r)]
25xLn((100+2*32,7)/(100-2*32,7))= 39.11 Cm.

CONCLUSIONS	
Rubric Score	**Lab Report fragment**

Insufficient *The conclusion is absent or does not refer to the research goal.*	**F21M12:** n chromosomes with Ler: 57, n chromosomes total: 116 - recombination frequency = 49.1 cM **Nga111:** n chromosomes with Ler: 32, n chromosomen total: 60 - recombination frequency = 53.3 cM **Ind2-40:** n chromosomes with Ler: 73, n chromosomes total: 142 - recombination frequency = 51.4 cM **T1B9:** n chromosomes with Ler: 72, n chromosomes total: 148 - recombinatiefrequentie = 48.6 cM **Ciw4:** n chromosomes with Ler: 26, n chromosomes total: 72 - recombination frequency = 36.1 cM **Ind4-6:** n chromosomes with Ler: 28, n chromosomes total: 58 - recombination frequency = 48.3 cM **Ind-4-12:** n chromosomes with Ler: 71, n chromosomes total: 126 - recombination frequency e = 56.3 cM **K18P6:** n chromosomes with Ler: 58, n chromosomes total: 136 - recombination frequency = 42.6 cM **Ciw9:** n chromosomes with Ler: 85, n chromosomes total: 152 - recombination frequency = 55.9 cM Thus, the closest marker is probably Ciw4.
Sufficient *The conclusion refers to the research goal but an explanation on the calculation of the recombination frequency is absent.*	The recombination frequencies of used markers are as follows: F21M12 - 33,3% Nga III - 19,4% Ind2-40 - 33,8% T1B9 - 29,2% Ciw4 - 35,9% Ind4-6 - 5,6% Ind4-12 - 45,9% K18P6 - 48,6% Ciw9 - 20,4% The probability that recombination occurs on a chromosome is bigger if the distance is also larger. Thus, the mutation will be near marker Ind4-6 since this one has the lowest recombination frequency. The mutation and the marker are thus close to each other, and the probability of a mutation is hence very low.
Excellent *The calculation of the recombination frequency is given and used to explain the most likely location of the gene of interest.*	The aim of this experiment is to locate the mutation causing the flowers to point downwards and the leaves to be directed with an angle of 90 degrees or less to the ground. It is also possible that the plants are smaller, but this is not always the case. This has been done with the aforesaid markers and results of the gel electrophorese. For every marker we detected if the F2 plants had a hit for Colombia, Landsberg or heterozygous. The locations of the markers were already known from the start, see figure 22. If we looked at our results we observed that a large part of our results was failed because the voltage we used for running the gel was too high.

Therefore, we could only use the results of the 2nd, 7 nd, 8th and 9th markers to fill in the form.

Figure 22, position of different markers on the chromosomes.



If the mutation is near a marker, there will be little crossing over between the mutation and the marker. This means that there is linked inheritance. If the marker is very far from the mutation, then the probability of recombination to occur is much larger. Hence it is possible to determine were the mutation is approximately located. You namely look for the marker where no recombination occurred. The marker with the lowest recombination is closest to the mutation. The degree of recombination per marker can be calculated with the recombination percentage (r). The formula to calculate r is as follows: C/(C+L) x 100%. The number of C and L can be calculated by looking at the number of hits with Colombia or Landsberg. A hit with a homozygous Colombia means that there is 2xC. The same accounts for Landsberg, only then you get 2xL. For a hit with a heterozygous you get 1xL or 1xC. The distance in centimorgan can then be determined with the following formula: D= 25 x ln[(100+2r)/(100-2r).

If we look at the distances, we see that for five markers it is not possible to calculate the distance. This is all right since these primers all had high recombination frequency, meaning that they are automatically far from the mutation. If we then look at the remaining values, we see that marker Ind4-12 has the smallest distance to the mutation. This means that our mutation is closest to the marker. A next step in the research could be to repeat the whole process and use even more markers near to Ind4-12. Like this, you can get closer and closer to the location of the mutation.

DISCUSSION	
Rubric Score	**Lab Report fragment**
Insufficient *The discussion is absent or does not discuss any limitations.*	From the results it would then become clear that the gene is located on chromosome 4, because this has the lowest recombination percentage. However, from the debriefing we know that the mutated gene is located on chromosome 1. That this does not match with the results is probably because not all experiments succeeded. This can result in deviations in the results that give another result.

Sufficient *Limitations such as the sample size and efficiency of PCR/gel electrophoresis are discussed.*	The mutant gene is present in every homozygous plant. In the experiment it was determined how often the wildtype marker was present. If the wildtype marker was not regularly present, an explanation would be that it is located near the mutant gene because it is difficult to be separated from the mutation. The least occurring recombination is the recombination at marker ind4-12. It is thus most probable that the mutant gene is, just like this marker, located on chromosome 4 near to this marker. A recombination frequency of 43, is however not very low and it is possible that it occurred by chance. An explanation for not finding a very low recombination percentage is that none of the markers is really close to ghe gene. This could explain why none of the recombination percentages is far below 50%. Another explanation is that the results are not very reliable because a very large part of the results of the gelelectrophoreses appeared not readable.
Excellent *Limitations are discussed and suggestions are made to use new markers for a more precise approximation of the exact location of the gene of interest.*	However, according to the debriefing of the experiment, the marker that is closest to the mutated gene is F12M12 on chromosome 1. This is not in accordance with the results of the whole group since this marker does not have the lowest recombination frequency. An explanation might be that the test is not reliable. Normally, there should be a reliable result with this many people. However, many of the gels were not readable and therefore it could not be determined whether recombination occurred. As a consequence, regardless of the high number of people and plants, a relatively low number of results were useful for drawing conclusions. The recombination frequencies are also less reliable because of the relatively low number of results. This explains why the results indicate that the marker closest to the gene is located on chromosome 5 instead of chromosome 1. Simply more F2 plants should be screened and as many results collected to determine whether the gene is indeed located on chromosome 1. This would result in a more reliable recombination frequency and will probably show that the mutated gene is indeed closest to the F21M12 marker on chromosome 1. Remarkably bands were not visible on some lanes on the gels. In a few cases bands were even not visible for all samples with the same primer pair. This is the case for primer pair F12, primer pair W4 and primer pair 4-6 (shown in figure 21 and 22). An explanation is that primers were not working, or no DNA was present. However, a sufficient amount of DNA was correctly isolated for all plants since DNA products are visible for other primer combinations. Thus, nothing went wrong during DNA isolation from plant tissue. There is also nothing wrong with Taq polymerase, since then you would expect to see no bands at all. Thus, the primer pairs seems to be the problem. An explanation would be that contamination with DNAses took place either in the mastermix or in the Eppendorf tube with primer pairs. The former group left the workplace very unclean. There was waste and plant soil everywhere. Of course, we removed the waste and tried to sterilize our workplace. However, the contamination of our workplace might explain why DNAses ended up in the Eppendorf tubes with mastermix or primer pairs. The DNAses will denaturate both primers and DNA in the 96 well PCR plate, with no DNA replication as a result.

DNA was also not visible in some other lanes, like primer nga and 2-40. Here it is also possible that DNA was not replicated because of the reason explained above. However, here the contamination with DNAses must have took place in the wells of the PCR plate itself, given that primers were not affected since bands of other plants are visible.