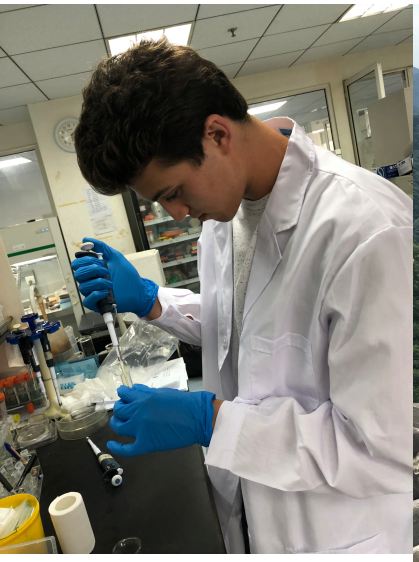
Identifying the Key Mechanisms Behind Successful Seedling Emergence in Rice



Chinese Academy of Agricultural Sciences
Beijing, China
2018 World Food Prize Borlaug-Ruan Intern
Rohan Freedman
Cape Elizabeth, Maine

Acknowledgments

Reflection:

Living in China was quite an adjustment from my life back in the United States. After overcoming jet lag, I had a range of other challenges. For example, one challenge I had was the language barrier. Luckily, many of the students in my lab were from other countries and spoke English well. However, most of the Chinese students spoke little to no English and I had to learn basic Chinese as well as master the art of using my hands to point and try to explain things. One thing I learned is that in the international community, English is the assumed common language. Everyone is expected to publish internationally in English as well as collaborate in English. Many times I was asked to help proofread papers from Chinese students who had tried to write in English. I was even offered a job to teach English. It made me realize, that although there are many languages in the world, if someone doesn't speak or understand English, we have isolated them and all their ideas. It made me realize how grateful I was to grow up speaking English. So, I decided to learn as much Chinese as possible so I could communicate with as many people in Beijing as I could. I tried to tear down the language barrier and converse with the farmer, the lady on the bus, the man at the pool representing the multitude of people whose stories I would never have heard if I only spoke English.

Another challenge I encountered was entering a PhD lab with no previous lab experience. All the machines and lab procedures were cutting-edge instruments that I had only read about in my AP Biology course. For example, I had studied all the steps involved in PCR (Polymerase Chain

Reaction) and in running a gel, but when I saw the machines I had no idea what buttons to press or even how to turn them on. I spent the first two weeks shadowing my lab mates and taking notes on common procedures such as isolating DNA. By the end of my eight weeks I used almost every machine in the lab and knew exactly what the character for “On” was in Chinese. This made me realize that despite studying AP Biology and its countless concepts, and stressing out over nearly 25 tests, I had learned less in one year than I had in that one day in the lab when I actually ran experiments. It was truly amazing being able to learn molecular biology through the real world application of the concepts I had learned in school. I am so grateful I had this opportunity.

Credits:

First, I would like to thank Dr. Norman Borlaug and Mr. John Ruan for founding the World Food Prize Foundation and creating this unbelievable opportunity for teenagers to learn among experts and join the global effort to combat food insecurity. In addition, I would like to thank all the amazing, dedicated people who believed in me and were involved in organizing such a unique internship. A special thanks to Ambassador Quinn and Crystal Harris, and all the staff at the World Food Prize Foundation. Without each of these people I would not have had such a remarkable experience as a Borlaug-Ryan intern. I will never be able to express how truly grateful I am.

Furthermore, I would like to thank the Chinese Academy of Agricultural Sciences for inviting me to live and learn at their school. I would like to extend my sincere gratitude to Dr.Li, my supervisor, at the Institute of Crop Sciences, along with my lab mates: Adeel, Suyash, Chun Yan, Uzair, Misan, Lu-Lee, Dr.Sunny, Chung Ying, Dr.Chow, Shang Chung Yuan, Mung Mung, and Dr.Fong. Thank you for all your mini lessons in molecular biology, help with Chinese words and lab procedures, and most of all, for being such friendly people. You are all such amazing intelligent people who will all change the world one day. I am so happy I got to meet all of you.

Finally, I would like to thank my World History II teacher, Ms.Page, who helped and encouraged me to apply for the Global Youth Institute as well as the Borlaug-Ruan internship. In addition, I want to thank all my friends and family back home who supported me throughout my adventure halfway around the world - especially my mom and dad, who provided me with endless inspiration, encouragement, and support. Thank you!

Abstract

For this study, we acquired a mini core collection of 300 lines of rice from around the world. Our purpose for conducting this study was to assess the natural variation of mesocotyl elongation in rice and its relation with seedling emergence. Mesocotyl elongation is an important trait that is not expressed in every line of rice. Our goal was to identify lines of rice with exceptionally long, average, and small mesocotyl lengths and correlate these with their respective GSK3 gene

sequences. Based on the results of our study we hoped to organize a select group of lines into haplotypes based on single nucleotide polymorphisms in the GSK3 gene.

Introduction

Rice is one of the world's most important crops accounting for a total global production of 480 million tons a year (Laporte, J., 2017). Billions of people rely on this staple crop to sustain life. However, as the world population continues to grow towards the projection of ten billion people by 2050, demand for rice will increase worldwide (World Population Projected, 2017, June 21). In addition, this discrepancy between supply and demand will be exacerbated by climate change which is destroying arable land and hindering global rice production.

The traditional method of growing rice requires a gargantuan amount of time, effort, and, most importantly, water. Many alternative methods have been explored and the direct seed sowing method has proven to be quite successful in saving time and water (Planting, 2015). However, there are many disadvantages to the mechanical seed sowing method such as weed competition, water availability, and seeds being sown too deep by mechanical planter (Planting, 2015). In this latter scenario, rice, has evolved a unique mechanism to overcome this problem by elongating the specialized part known as the mesocotyl. The mesocotyl is a part of the rice seedling that elongates to push the seedling up allowing the plant to emerge from the soil. *See Figure 1 for more information (in the supporting information section) on mesocotyls and rice seedling anatomy.* This life saving trait has been observed and studied, however, its specific molecular

mechanisms are not completely understood. With a changing climate, methods for developing rice that can grow in adverse conditions are necessary. Assessing mesocotyl elongation is important because deep sowing is advantageous and the mesocotyl plays a crucial part in the success of deep sowing. Better defining the mechanisms behind mesocotyl elongation will maximize the benefits that come from deep sowing.

Prior to this experiment, we believed that we would observe genotypic variability between the different mesocotyl elongations and that we could associate specific nucleotide differences to account for the variations.

Method

Participants:

Our experiment included 300 genetically characterized rice lines from over 50 different countries drawn from every continent except Antarctica. This mini core collection was designed to be representative of the entire global cultivated rice population. Each strain was adapted to its specific environment and was significant in allowing us to analyze the variation amongst rice grown in adverse climates.

Apparatus and Materials:

<u>Preparing growth tubes and sowing seeds:</u>

• Transparent glass bottles - 500 ml

• Vermiculite soil

- 30 seeds
- 200 ml of distilled water at pH 5.8
- Incubator

Growing the seedlings:

- Cardboard boxes
- Black polythene bags to cover bottles and simulate darkness
- Controlled temperate room for storage

Measuring the seedlings:

- Canon EOS camera
- 75% Ethanol
- Image J software
- 20 cm ruler for scale

DNA isolation:

- Warm water bath
- 1.5 ml tubes
- CTAB Buffer
- Trichloromethane
- Ethanol
- Pipettes
- Table top centrifuge
- Liquid nitrogen
- Vertical shaker with 3mm diameter medal beads

PCR preparation:

- Table top centrifuge
- Pipettes
- dNTP
- DNA
- PCR strip tubes
- Buffers - kept in -20ºC freezer
- Enzymes - kept in -20º C freezer
- RNA primers (TsingKe Biological Technology) - F and R primer
- PCR machine

Gel electrophoresis:

- DNA
- Table top centrifuge
- Loading dye
- Agarose Gel
- TAE Buffer
- Ethidium Bromide
- Nanodrop spectrophotometer

Sequencing Company:

- TsingKe Biological Technology

Data storage:

• SnapGeneViewer

• Excel spreadsheet

Procedure:

To begin, 30 seeds from each of the 300 lines were put in little parchment seed bags. These bags were put into an incubator at 45ºC for 24 hours so the seeds would leave dormancy. After this, 500ml glass bottles were gathered and each was filled with vermiculite soil 5 cm from the bottom. Then each bottle was labeled with the corresponding line number and 30 seeds were sowed evenly into the vermiculite. Additional vermiculite soil was then added to each bottle until it covered the seeds and was 2 cm from the top of the bottle. Then 200 ml of distilled water was added to each bottle at a pH of 5.8 (adding NaOH and HCl as needed to achieve 5.8 pH). Finally, the 300 bottles were placed in six large cardboard boxes, covered in black bags, then sealed shut with duct tape. Each bottle was kept in complete darkness - this is important because mesocotyl elongation is light sensitive and elongation occurs only in darkness until the seed finds light and senses that it has emerged. *See Figure 2 for a comparison of light grown v.s dark grown seedlings.* The boxes were moved into a controlled climate room set at 28ºC. The seeds were left to grow in darkness for ten days.

After ten days, the bottles were cleaned out and the ten-day old rice seedlings were washed and then put back in their bottles which was now filled with ethanol - to preserve the seedling and

make sure it stopped growing. *See Figure 3 for an image of a bottle after 10 days.* Seedling germination was variable; there were lines with anywhere from 5 to 30 seedlings that germinated. Following this, each seedling within a bottle was taken out and lined up on a black surface with a 20 cm ruler for scale. Pictures were taken for each bottle and the pictures were analyzed on a software called ImageJ. ImageJ was used to trace over the part of the seedling defined as the mesocotyl and calculate the measurement based on the known pixels per inch from the ruler. This data provided us with approximately fifteen measurements of mesocotyl length for each of the 300 lines. This data was entered into an Excel spreadsheet and the averages for each line were calculated.

Following this, 15 lines were chosen as samples used for further analysis. The 15 lines included: 5 lines with the longest mesocotyl, 5 lines with the most average mesocotyl length, and 5 lines with the smallest to no mesocotyl observed. Five seeds from each of the 15 lines were regrown using the same previous method and conditions. After they germinated, the DNA of each was isolated. To carry this out, a part of the shoot from each rice seedling was cut and put in a 2 ml tube along with a 2 mm metal ball. The tubes were put in liquid nitrogen and then put on a vertical shaker that crushed the leaf samples. Then 700 µl of CTAB buffer was added to the tube, (to lyse the plant cell) and the samples were put in a 65ºC water bath for 45 mins. After incubation, 700µl of tricholormethane was added to each tube as a solvent. The samples were then centrifuged at 12000 rpm for ten minutes. This separated the plant cell and cell wall. The DNA was in the liquid part of the sample and 500 ml of it were extracted using a pipette and put

into 1.5 ml tubes. Then ethanol was added to precipitate the DNA. The samples were then moved to a 4ºC fridge for ten minutes and then centrifuged at 12000 rpm for ten minutes. This separated the DNA (which precipitated at the bottom) from the excess material that was in the liquid. The liquid was thrown away and 1ml of ethanol was added again to clean the sample and the samples were centrifuged once again. Then, the ethanol was removed by pipettes and 50 µl of distilled water was added. The tubes now each held the genomic DNA from each of the 15 lines.

Using the 50 µl of diluted DNA, a gel electrophoresis was conducted to check if the DNA had been isolated and was ready for PCR. Using distilled water, TAE buffer, ethidium bromide, and agarose, a gel was created. 3 µl of DNA was mixed with 0.3 µl of DNA marker and the 15 lines of DNA were each put in a separate well. The machine was turned on and the negative charge of the DNA was attracted to the positive end of the gel and it moved through the agarose gel. We cut the gel with the 15 wells shown and put the gel in a Trans UV machine that uploaded images onto a computer for analysis. The ethidium bromide is a fluorescent chemical that binds to DNA and allows us to visualize the DNA under UV light. This confirmed we had successfully isolated DNA for most of the 15 lines. *See Figure 4 for the gel results.* In addition, 1 µl of each line's DNA was put on a Nanodrop Spectrophotometer which used known wavelengths (260 -280 nanometers) to quantify the concentration of DNA in our samples.

Following this, the sample was prepped for PCR (Polymerase Chain Reaction). PCR is a technique used to amplify target DNA to ultimately generate millions of copies of a particular DNA sequence. PCR works by altering the chamber temperature to reach high temperatures that denature the DNA into its two strands and then lowers the temperature for the annealing and extension phase where complementary nucleotides attach to the two strands. This creates two identical double strands of DNA from the single original DNA. This cycle is repeated many times to exponentially increase the number of DNA copies.

To run PCR, 3 µl of sample DNA from each line was added to a master mix containing 25 µl of a buffer, 16 µl of distilled water, 1µl of polymerase enzyme, two forward RNA primers and two reverse RNA primers for isolating the GSK3 gene (designed in the lab and ordered from TsingKe Biological Technology), and four dNTPs. The 15 samples were then moved into PCR strip tubes and put into the PCR machine. The PCR machine was set to 36 cycles with a denaturing temperature of 97ºC, annealing temperature of 56ºC, extension temperature of 72ºC, and a hold temperature of 4ºC. The PCR took around four hours. After the PCR was complete, a sample of each line was taken and run on a gel to check if the concentration of the target DNA had increased. The extremely bright lines seen on a gel were clear evidence of a successful PCR - the Nanodrop Spectrophotometer confirmed this. *See Figure 5 for the gel image.*

Now that the GSK3 gene had been successfully isolated by primers and multiplied in concentration by PCR, the samples were ready to be sequenced. After packaging the 15 samples of the GSK3 gene, a form was filled out for TsingKe Biological Technology and the sample was

stored in the 4ºC fridge ready for pickup. The sequencing guy from TsingKe Biological Technology came to the lab and picked up our samples and 24 hours later a email was sent with the sequencing results. The 15 gene sequences were stored and analyzed in SnapGene Viewer. The sequencing results we had were from genomic DNA and not cDNA. This meant the DNA still had introns that needed to be spliced. Despite this, the coding regions and introns were identified with the help of the Michigan State University Rice Genome Annotation project.

At this point, the 15 sequencing results of the GSK3 gene from each line were compared and specific nucleotide differences in specific places were found between the 15 lines. The four identified SNPs helped categorize the 15 lines into either haplotype Type A or haplotype Type B. The assigned haplotype were then analyzed to see if a correlation existed between haplotype and mesocotyl length. More specifically, if haplotype Type B correlated with longer mesocotyl length. *The results are shown in Figure 6.*

Finally, the results of the haplotyping and correlation study were cross checked with a *Nature* Publication. The *Nature* publication ran a similar study and used the gene OsGSK2 - a *GSK3*-like kinase (Sun, S et al. 2018, June 28). *See Figure 7 for the Nature publication haplotyping results.*

Results and Discussion

The goal of this study was to better define the mechanisms behind mesocotyl elongation in rice seedling development. The hypothesis was that a significant correlation between specific SNPs in the GSK3 gene and mesocotyl length would exist and be observed in this study. The average mesocotyl for each line was calculated and each line was sequenced. SNPs were observed at the 213th, 390th, 394th, and 401st codon. Type A had ACA, CGA, ACC, and GGG at each position, respectively. Type B had GCA, CAA, GCC, and AGG at these same positions. These SNPs change the codons and ultimately the protein synthesized. Consequently, Type A had the amino acids: threonine, arginine, threonine, and glycine produced at the 213th, 390th, 394th, and 401st positions where the SNPs in Type B haplotype coded for the amino acids : alanine, glutamine, alanine, and arginine (The Single-Letter Amino Acid Code [PDF]). *See Figure 8 and 9 for a summary of these results.*

It is proposed that these amino acid differences account for a significant change in protein shape and function resulting in longer mesoctyls. However, the results of this study showed a partial trend, but not complete correlation of haplotype with mesocotyl length, suggesting that perhaps more than one genetic locus is involved in determination of mesocotyl length. Lines 1, 3, 9, 10, and 13 were the only lines classified as Type B, however, they were not all part of the longest mesocotyl group (two were part of the longest group and three were part of the medium group). None of the Type B lines were in the short to no elongation group. *See Figure 6 for a summary of the lines and grouping.*

The results of this study provide evidence that supports the idea that there may be some unknown genes which regulate this trait in addition to the GSK3 gene. This would suggest that the next step in this study would be to conduct a Genomic Wide Association Study (GWAS) to search for other possible genes responsible for this trait.

It is important to note that due to time constraints, this experiment was run only once. Regardless, it is generally rare to find phenotypes associated with a single gene. Most traits in nature and plants are linked to more than one gene. According to this study and its results, GSK3 may not be the only gene responsible for the expression of mesocotyl elongation. Discovering the other, as yet unknown, gene(s) will be crucial in defining the mechanisms behind this life saving trait in rice seedlings. If the genes responsible for this trait as well as the associated SNPs are discovered, then farmers could select these resilient lines, potentially yielding a higher success rate for seed germination, plant development, and most importantly, higher total crop yield. In conclusion, further defining the key mechanisms behind this trait has the potential to increase global rice production to ultimately help feed the ten billion people that are projected to inhabit our planet in 2050.

References

- Laporte, J. (2017). Topic: Rice. Retrieved September 30, 2018, from https:/ www.statista.com/topics/1443/rice/

- Planting. (2015). Retrieved October 1, 2018, from http://www.knowledgebank.irri.org/ ericeproduction/II.3_Direct_seeding.htm

- Rice Genome Annotation Project. (2013, February 6). Retrieved September 30, 2018, from http://rice.plantbiology.msu.edu/

- Sun, S et al. (2018, June 28). Natural selection of a GSK3 determines rice mesocotyl domestication by coordinating strigolactone and brassinosteroid signaling. Retrieved September 30, 2018, from https://www.nature.com/articles/ s41467-018-04952-9

- The Single-Letter Amino Acid Code [PDF]. Retrieved from http://130.88.97.239/ bioactivity/aacodefrm.html

- World Population Projected. (2017, June 21).| UN DESA Department of Economic and Social Affairs. Retrieved from https://www.un.org/development/desa/en/news/ population/world-population-prospects-2017.html

Supporting Information

Figure 1

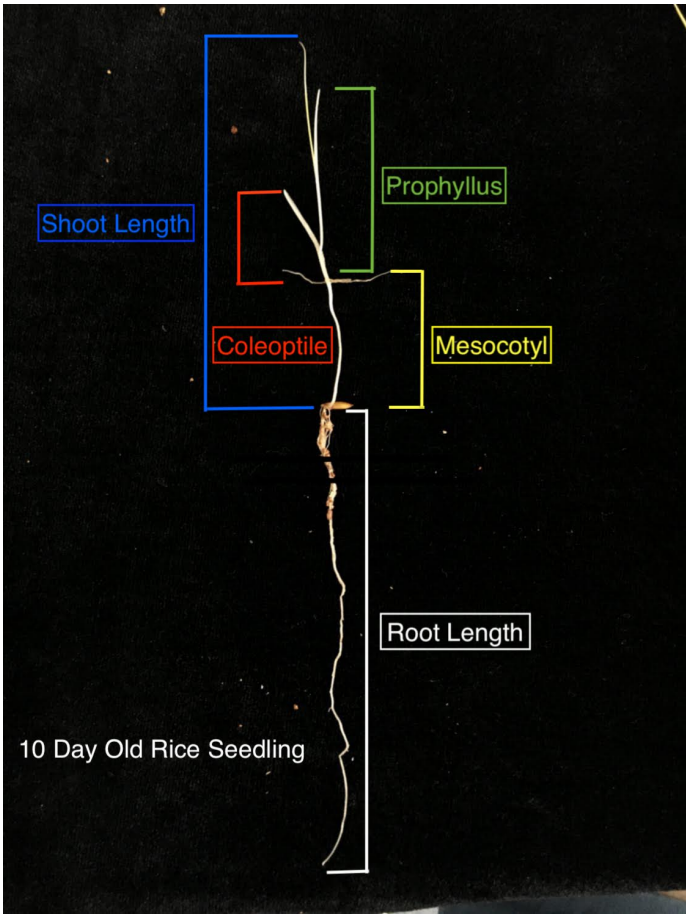


Figure 1 shows the anatomy of a 10 day old rice seedling. The mesocotyl is highlighted in yellow. Notice that the mesocotyl is the tissue that elongates from the seed and pushes the coleoptile and prophyllus up.

Figure 2

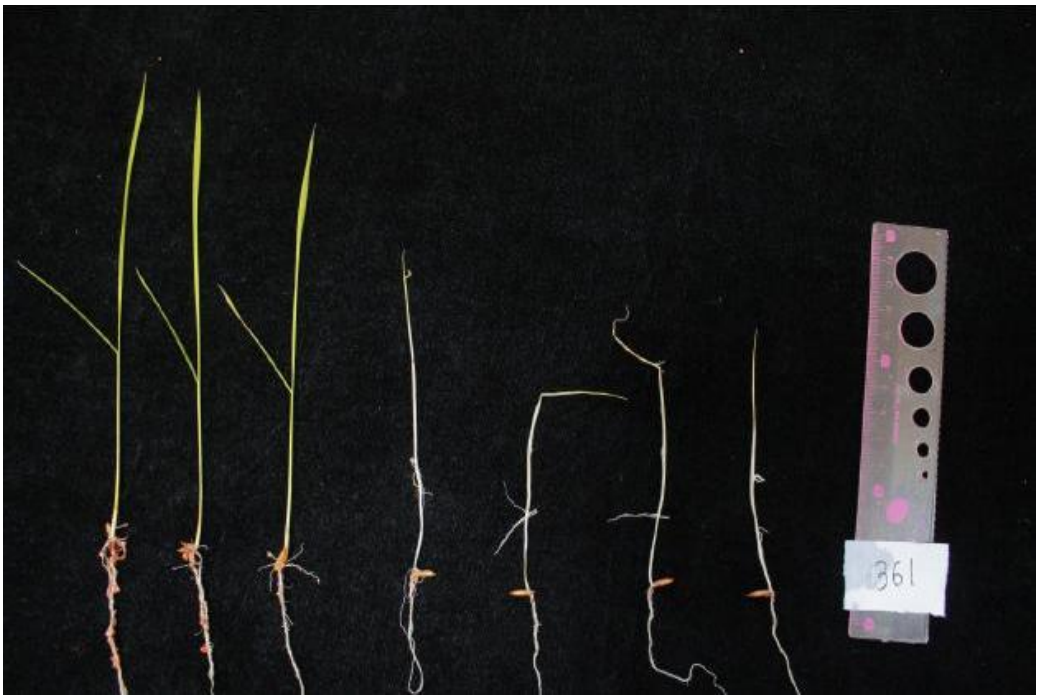


Figure 2 shows 7, 10 day old rice seedlings. The 3 on the left were exposed to light while the 4 on the right were not. It is important to not that the 4 on the right have elongated mesocotyls while the 3 on the left do not. Also, the 4 on the right are lacking pigment due to the lack of sunlight and chlorophyll.

Figure 3

Figure 3 shows one of the bottles that grew a rice seedling in dark for 10 days. This line seems to have had successful seed emergence given that so many seedlings emerged from the soil.

Figure 4

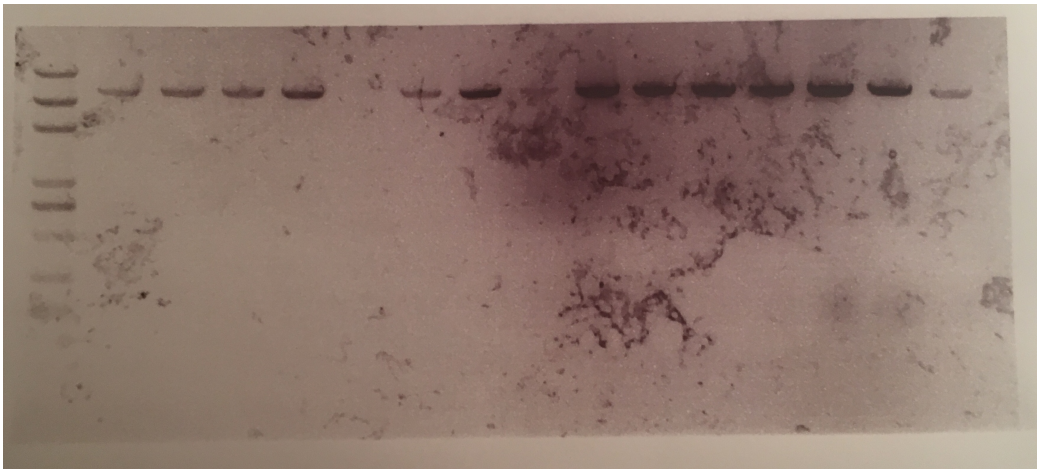


Figure 4 shows the results of a gel that ran the isolated DNA from the 15 selected lines.

Figure 5

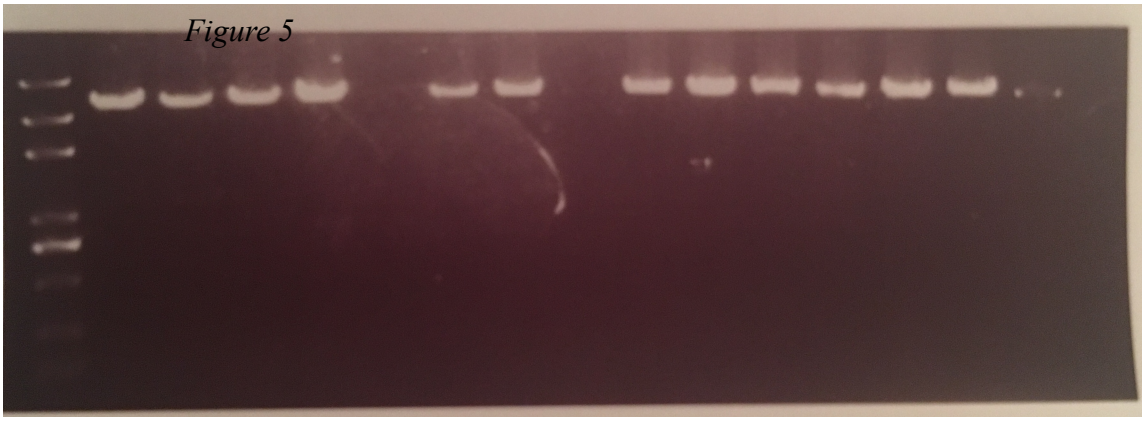


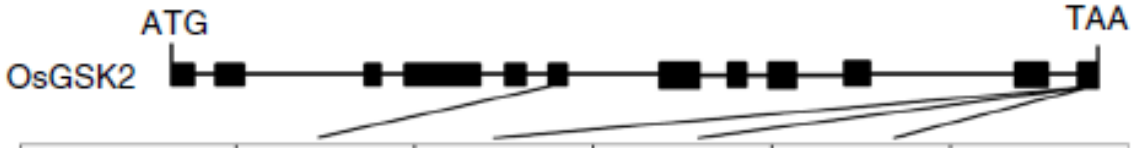
Figure 5 shows the results of a gel that ran the amplified GSK3 gene for the 15 selected lines. It is important to notice that the bands on this gel are brighter here than in figure 4 because PCR substantially increased the concentration of DNA in each sample.

Figure 6

Results of Selected 15 Lines of Rice

Line Number	Average Mesocotyl Length (cm)	Name	Country	Haplotype
1	1.734	KEERIPALA CHILL PADDY::IRGC 49790-1	India	Type B
2	0.1	XH211	Unknown	Type A
3	4.75	M 142::IRGC 35054-1	India	Type B
4	4.03	KALU T 139::IRGC 53670-1	Bangladesh	Type A
5	6.906	3210::IRGC 116950-1	Thailand	Type A
6	2.4099	Basmati 385	Pakistan	Type A
7	0.1	HP 3319-2WX-6-4-1-B::IRGC 117331-	Republic of Korea	Type A
8	1.3	Xiangwanxian 1	China	Type A
9	1.0866	NERICA 8	Cote d'Ivoire	Type B
10	2.917	EX FOILAEIN (NAPUTO)::IRGC 81675-1	Nigeria	Type B
11	0.1	HSINCHU AI CHIO CHIENG::IRGC 78182-1 12 94 0.03 SAN ZHAO QI::IRGC 72289-1	Taiwan	Type A
12	0.03	SAN ZHAO QI::IRGC 72289-1	China	Type A
13	6.1	UCP 122::IRGC 8794-1	Bangladesh	Type A
14	0.1	XH210	Unknown	Type A
15	3.98	BIR BAHADUR::IRGC 53889-1	India	Type A

Figure 7 **Osgsk2 Natural Variation and Haplotypes**



No. of accessions	SNP1	SNP2	SNP3	SNP4	Genotype
313	T (T^{213})	C (R^{390})	T (T^{394})	C (G^{401})	Type A
190	C (A^{213})	T (Q^{390})	C (A^{394})	T (R^{401})	Type B
1	T (T^{213})	C (R^{390})	T (T^{394})	T (R^{401})	Type C
1	C (A^{213})	C (R^{390})	C (A^{394})	C (G^{401})	Type D

NATURE COMMUNICATIONS | (2018)9:2523 | DOI: 10.1038/s41467-018-04952-9 | www.nature.com/naturecommunications

Figure 8 **Sequence Analysis of the Selected Lines**

Figure 9

Single Nucleotide Polymorphism(SNP) grouping into Two

Line Number	SNP1	SNP2	SNP3	SNP4	SNP GROUP
Reference	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
1	GCA (A213)	CAA (Q390)	GCC (A394)	AGG (R401)	Type B
2	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
3	GCA (A213)	CAA (Q390)	GCC (A394)	AGG (R401)	Type B
4	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
5	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
6	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
7	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
8	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
9	GCA (A213)	CAA (Q390)	GCC (A394)	AGG (R401)	Type B
10	GCA (A213)	CAA (Q390)	GCC (A394)	AGG (R401)	Type B
11	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
12	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
13	GCA (A213)	CAA (Q390)	GCC (A394)	AGG (R401)	Type B
14	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A
15	ACA (T213)	CGA (R390)	ACC (T394)	GGG (G401)	Type A