Predicting Functional Impacts of Amino Acid Substitutions in Exon 21 of the ATP7B Gene for Wilson Disease Diagnosis

Omar Qahtan Yaseen [1*], Asra'a Adnan Abdul-Jalil [2*]

Abstract

Background: Diagnosing illnesses with overlapping clinical symptoms shows challenges, necessitating precise identification of genetic variations underlying pathogenesis. Here, we focus on Wilson disease, an autosomal recessive disorder characterized by copper accumulation due to mutations in the ATP7B gene. Method: To predict the functional impact of single amino acid substitutions (SAASs) in exon 21 of ATP7B, we employed bioinformatics tools, including SIFT, PolyPhen2, and Provean. Our study, conducted on thirty Iraqi Wilson disease patients, identified missense mutations associated with disease manifestation. Result: Bioinformatics analyses revealed nine potentially deleterious non-synonymous SNPs in exon 21. Functional modifications were predicted more accurately by all programs, indicating their utility in identifying pathogenic variants. Conclusion: Our findings underscore the utility of computational methods in high-throughput SAAS annotation, offering insights for diagnostic screening and therapeutic strategies. Furthermore, our study expands the spectrum of ATP7B mutations implicated in Wilson disease onset, underscoring the role of bioinformatics in elucidating genotype-phenotype correlations and advancing precision medicine.

Keywords: SAASs, ATP7B gene, Wilson disease, Bioinformatics algorithms, Missense mutations.

Significance | This study determined the SAASs' impact on protein function aids in diagnosing genetic diseases like Wilson disease using bioinformatics algorithms.

*Correspondence. Omar Qahtan Yaseen, Department of Heet Education, General Directorate of Education in Anbar, Ministry of Education, Hit, Anbar 31007, Iraq. And Asra'a Adnan Abdul-Jalil College of Pharmacy, University of Anbar, Iraq. E-mail: omerkahtan@gmail.com, sc.dr_asraa2017@uoanbar.edu.iq

Editor Md Shamsuddin Sultan Khan, And accepted by the Editorial Board Mar 31, 2024 (received for review Feb 6, 2024)

Introduction

Differentiating between illnesses that share similar clinical symptoms can be challenging (Kandil et al., 2019; Wang et al., 2012). Certain single amino acid substitutions (SAASs), referred to as functional SAASs, can disrupt regular protein function, leading to evident changes in morphology or physiological state (Li et al., 2012; Wang et al., 2016; Xu et al., 2018). SAASs are the most prevalent variations, according to extensive variety investigations of human genomes, including those from malignant tumors (Lek et al., 2016). It is challenging to separate substantial effect alterations from these other variant versions and functional differences from all the variants due to the large number of single amino acid substitutions dispersed across the genome. Nevertheless, determining how SAASs affect proteins provides knowledge about the molecular processes driving complicated characteristics and the functional areas and interactions of genes and proteins. It is also vital to annotate these linkages (Kono et al., 2018; Kovalev et al., 2018). Conventional experiments can precisely determine how SAASs affect protein function, but they are labor-intensive and costly in terms of resources, being difficult to modify, and taking time (Xu et al., 2023).

Author Affiliation.
[1] Department of Heet Education, General Directorate of Education in Anbar, Ministry of Education, Hit, Anbar 31007, Iraq.
[2] College of Pharmacy, University of Anbar, Iraq.

Please cite this article.
Omar Qahtan Yaseen, Asra'a Adnan Abdul-Jalil. (2024). Predicting Functional Impacts of Amino Acid Substitutions in Exon 21 of the ATP7B Gene for Wilson Disease Diagnosis, Journal of Angiotherapy, 8(3), 1-6, 9643

Moreover, the accumulation of data from whole-genome sequencing and resequencing analysis in initiatives such as the new assembly for a pan-genome in plants and animals (Zhao et al., 2018; Hufford et al., 2021; Tao et al., 2021) has resulted in a notable rise in the amount of SAAS, making these traditional approaches ineffective. The prediction of SAASs' impact on protein structure and function using computational methods, identifying functioning SAASs for experimental purposes investigation, is one possibility for high-throughput SAAS annotation (Kovalev et al., 2018; Xu et al., 2023). Copper builds up in the liver, brain, and eyes in Wilson disease, an autosomal recessive disorder. Serious brain damage or the requirement for long-term medical attention might arise from the disease's early beginning. Another significant consequence of the illness that may necessitate a liver transplant is the progression of acute liver failure and liver failure. Compared to other cultures, East Asian people have a significantly greater prevalence of Wilson illness, which typically ranges from 1 in 30,000 to 1 in 50,000 (Kok et al., 2008; Gavali et al., 2024). The condition is caused by mutations in the ATP7B gene, which is found on chromosome 13 at 13q14.3. This gene codes for a P-type copper-transporting adenosine triphosphatase (Kok et al., 2008). Detecting ATP7B gene mutations is a useful method for Wilson disease diagnosis and treatment strategy. Several software tools have been created to forecast the impact of SAASs on human protein function (Ioannidis et al., 2016; Alirezaie et al., 2018; Chennen et al., 2020; Pejaver et al., 2020; Takeda et al., 2020). Gene changes associated with human diseases are often identified using family-based linkage analysis or population-based association studies. A linkage analysis looks for genetic markers that are co-inherited with a query trait (usually ranging in length from 1 to 5 million bp) in order to identify sensitive disease-causing loci. In complicated disorders where several vulnerable gene alterations interact with environmental variables, linkage analysis performs poorly and has little predictive power for hard-to-collect family-based sequencing data (Verkade et al., 2016). We identified ATP7B gene mutations in several Wilson disease-affected youngsters using Sanger sequencing techniques. To estimate how mutations will affect pathogenesis, we employed prediction software like SIFT, PolyPhen2, and Provean.

Materials and methods

Thirty Wilson Disease patients, aged three to fourteen, including six females and twenty males, were recruited from western Iraq. Genomic DNA was extracted from peripheral blood samples. We looked for mutations in the ATP7B gene's exon 21. All research activities described in this paper were conducted in accordance with ethical guidelines and approved by the University of Anbar institutional review board.

DNA extraction

DNA extraction was performed using a commercially available kit (Qahtan, 2023).

Primer

The primer for exon 21 of the ATP7B gene was amplified using the information in the Table 1 (Al-Mayah, Q. S., et al 2016).

Polymerase chain reaction (PCR)

Polymerase Chain Reaction (PCR) amplification was conducted according to the protocol described by Qahtan (2023).

Bioinformatics algorithms

Bioinformatics algorithms were employed to predict the functional impact of detected variations on disease pathogenesis (Table 2).

Results and discussion

Protein function relies on local and global dynamics. Disease-causing mutations often destabilize proteins, weaken binding, and increase flexibility. Our study broadened the spectrum of ATP7B mutations associated with Wilson Disease. Bioinformatics screening can detect ATP7B mutations and assess SNP effects.

Wilson disease, an uncommon hepatobiliary illness prevalent among Asian youngsters, particularly in Iraq, is often linked to missense mutations in exon 21 of the ATP7B gene, identified through genetic testing (Qahtan, O. 2023). Genetic, endogenous, and environmental factors can influence an individual's susceptibility to oxidative damage due to inherited mutations. Single nucleotide polymorphisms (SNPs) are implicated in various genetic disorders, including Wilson disease, highlighting the significance of genetic variations (Al-Mayah, Q. S., et al 2016; Gul, B., et al 2022). Over 4 million distinct SNPs have been identified, offering insights into the genotype-phenotype relationship (http://www.ncbi.nlm.nih.gov/SNP/index.html). However, mechanisms underlying the phenotypic effects of SNPs remain poorly understood.

The SIFT program classified variations (V20G, R41E, Q23H, and N36K) as tolerant, while polyphen2 and provean predicted functional impact (Figure 1, Figure 2). Conversely, SIFT's intolerant variations (I31S, S13L) showed no effect in provean and polyphen2. In exon 21 of ATP7B gene, nine potentially deleterious nsSNPs were identified using in silico techniques (Table 3). SIFT predicts phenotypic impact and distinguishes functional from non-functional polymorphisms based on sequence conservation and physiochemical properties (Figure 1, Figure 2). Polymorphism Phenotyping v2 utilizes physical and comparative factors to forecast amino acid mutation effects on human protein structure and function. Identification of these nsSNPs may aid in developing rapid and cost-effective screening methods for ATP7B-related illnesses.

Diagnostic failure is a primary cause of mortality in Wilson Disease (WD), underscoring the urgency of prompt diagnosis and treatment, even in the presence of hepatic comorbidities. SIFT

Table 1. primer used in this study

Forward	5'-CCCAACTTGTGTAGCTGCTG - 3'
Reverse	5'-GGCCAACTGGTGCTTACTTT- 3

Table 2. Bioinformatics tools in this study

Tool	Web
SIFT	(https://sift.bii.astar.edu.sg/),
PolyPhen2	(http://genetics.bwh.harvard.edu/pph2/),
PROVEAN	(http://provean.jcvi.org/seq_submit.php).

Table 3. nsSNPs predicted by *Insilco* programs

No. Of sample	Nucleotide change	Variant	Predicted Effect	SIFT score	Prediction	Provean score	Prediction	Polyphen-2 score	Prediction
1	T>G	V20G	Missense	1.00.	Tolerated	-6.970	Deleterious	1.000	Probably Damaging
	A>G	R41E	Missense	1.00.	Tolerated	-4.191	Deleterious	0.895	Possibly Damaging
2	G>C	Q23H	Missense	0.73	Tolerated	-2.781	Deleterious	0.783	Possibly Damaging
	C>A	N36K	Missense	0.14	Tolerated	-6.000	Deleterious	1.000	Probably Damaging
	C>A	P43T	Missense	1.00	Tolerated	0.083	Neutral	1.000	Probably Damaging
3	A>C	N29P	Missense	1.00	Tolerated	-0.167	Neutral	0.009	Benign
4	T>G	I31S	Missense	0.00	Deleterious	-0.061	Neutral	0.002	Benign
5	A>G	S45G	Missense	0.28	Tolerated	-0.633	Neutral	0.045	Benign
6	T>C	L18P	Missense	0.28	Tolerated	-7.000	Deleterious	0.000	Benign
7	A>T	S13L	Missense	0.00	Deleterious	-0.444	Neutral	0.134	Benign
8	A>G	K13E	Missense	1.00	Tolerated	-3.000	Deleterious	0.048	Benign

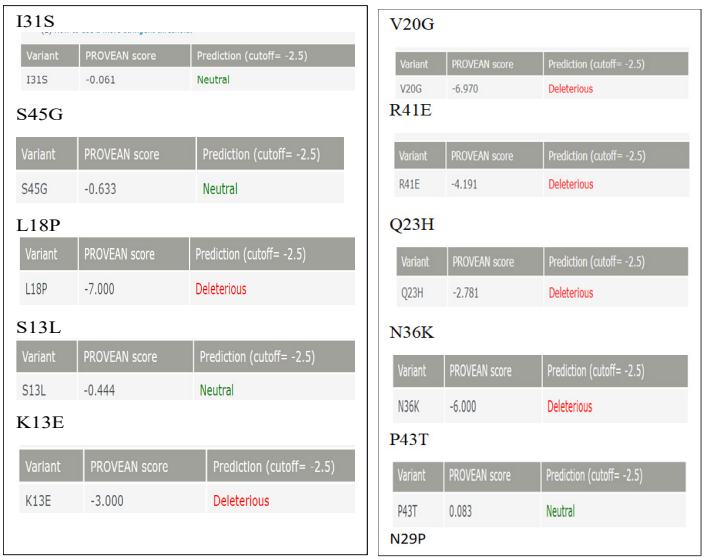
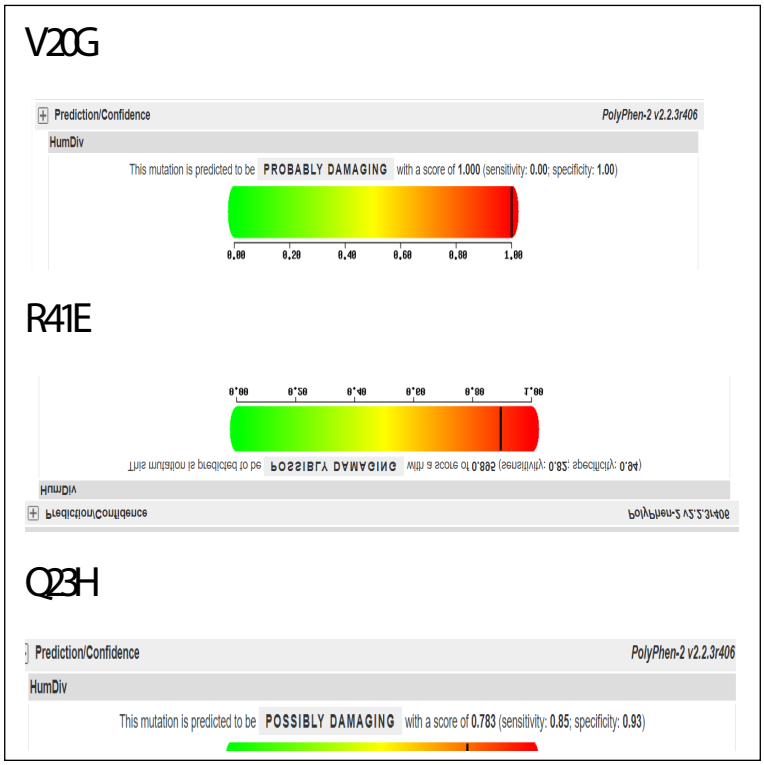


Figure 1. Prediction of the functional impact of detected variations on disease pathogenesis Provean tool.



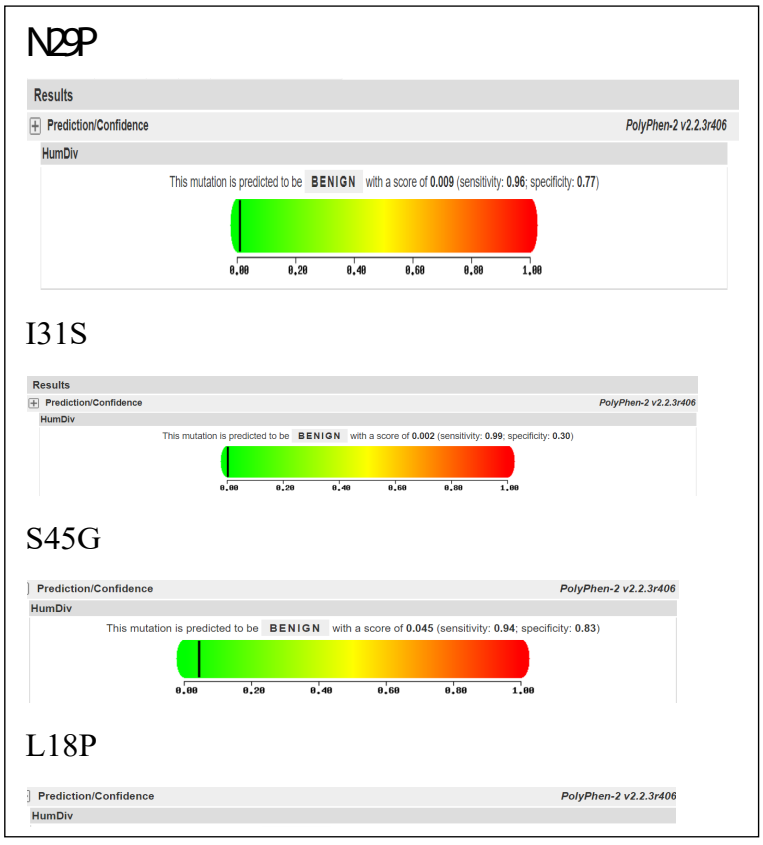


Figure 2. Prediction of the functional impact of detected variations on disease pathogenesis polyphen2 tool.

calculates the likelihood of an amino acid substitution negatively impacting protein function based on sequence homology. This premise suggests that evolutionarily conserved regions are less mutagenic, making changes more likely to affect function (Maghool, F., et al., 2023; Yaseen, O. Q., et al., 2020; Tasmeen, R., et al., 2022; Xiong, D., Lee, et al., 2022; Lei, P., et al., 2021).

Conclusions

In conclusion, our study underscores the utility of computational methods in predicting the functional impact of single amino acid substitutions (SAASs) in the ATP7B gene, particularly in exon 21, which is crucial in Wilson Disease (WD) diagnosis. By employing bioinformatics tools such as SIFT, PolyPhen2, and Provean, we identified potentially deleterious nsSNPs associated with WD manifestation. This expands the spectrum of ATP7B mutations implicated in WD onset, aiding in genotype-phenotype correlations and advancing precision medicine. Prompt diagnosis and treatment are imperative, given the diagnostic failure's significant contribution to WD mortality. Our findings highlight the role of bioinformatics in diagnostic screening and therapeutic strategies, offering hope for improved patient outcomes.

Author contributions

O.Q.Y. and A.A.A.J. designed the study, wrote the manuscript, read and approved the manuscript.

Acknowledgment

Authors were grateful to each research participant and their guardians. Additionally, I thank Dr. Yassin H. Majeed, a consultant in internal medicine and liver diseases, and a professor in the Department of Internal Medicine and Gastroenterology at the University of Anbar's College of Medicine in Ramadi City, Anbar, Iraq, for his thoughtful assistance and hard work.

Competing financial interests

The authors have no conflict of interest.

References

Alirezaie, N., Kernohan, K.D., Hartley, T., Majewski, J. and Hocking, T.D.(2018) ClinPred: prediction tool to identify disease-relevant nonsynonymous singlenucleotide variants. Am. J. Hum. Genet. 103, 474–483.

Al-Mayah, Q. S., Arif, H. S., Shamran, H. A., Abbas, A. A. Q., & Kareem, K. H. (2016). Mutation Profile of Exon 14 and Exon 21 of ATP7B Gene in Patients with Wilson Disease. Medical Journal of Babylon, 13(3), 635-641.

Chennen, K., Weber, T., Lornage, X., Kress, A., Bohm, J., Thompson, J., ¨ Laporte, J. et al. (2020) MISTIC: A prediction tool to reveal disease-relevant deleterious missense variants. PLoS One, 15, e0236962.

Gavali, L. V., Mohammed, A. A., Al-Ogaili, M. J., Gaikwad, S. H., Kulkarni, M., Das, R., & Ubale, P. A. (2024). Novel terephthalaldehyde bis (thiosemicarbazone) Schiff base ligand and its transition metal complexes as antibacterial Agents: Synthesis, characterization and biological investigations. Results in Chemistry, 7, 101316.

Gul, B., Firasat, S., Tehreem, R., Shan, T., & Afshan, K. (2022). Analysis of Wilson disease mutations in copper binding domain of ATP7B gene. Plos one, 17(6), e0269833.

Hufford, M.B., Seetharam, A.S., Woodhouse, M.R., Chougule, K.M., Ou, S., Liu, J., Ricci, W.A. et al. (2021) De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. Science, 373, 655–662.

Ioannidis, N.M., Rothstein, J.H., Pejaver, V., Middha, S., McDonnell, S.K., Baheti, S., Musolf, A. et al. (2016) REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am. J. Hum. Genet. 99, 877–885.

Kandil, M. A., Aiad, H. A., El-Azab, D. S., El-Goday, S. F., & Tantawy, M. S. (2019). Clinicopathological differentiation between biliary atresia and other causes of neonatal cholestasis. Menoufia Medical Journal, 32(1), 151.

Kok, K. F., Hoevenaars, B. M., Waanders, E., & Drenth, J. P. H. (2008). Value of molecular analysis of Wilson's disease in the absence of tissue copper deposits: a novel ATP7B mutation in an adult patient.

Kono, T.J.Y., Lei, L., Shih, C.H., Hoffman, P.J., Morrell, P.L. and Fay, J.C.(2018) Comparative genomics approaches accurately predict deleterious variants in plants. G3, 8, 3321–3329.

Kovalev, M.S., Igolkina, A.A., Samsonova, M.G. and Nuzhdin, S.V.(2018) A pipeline for classifying deleterious coding mutations in agricultural plants. Front. Plant Sci. 9, 1734.

Lei, P., Ayton, S., & Bush, A. I. (2021). The essential elements of Alzheimer's disease. Journal of Biological Chemistry, 296.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H. et al. (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature, 536, 285–291.

Li, B.B., Zhao, Y.X., Liang, L.Y., Ren, H.B., Xing, Y., Chen, L., Sun, M.Z. et al. (2012) Purification and characterization of ZmRIP1, a novel reductantinhibited protein tyrosine phosphatase from maize. Plant Physiol. 159, 671– 681.

Maghool, F., Emami, M. H., Alipour, R., Mohammadzadeh, S., Sereshki, N., Dehkordi, S. A. E., ... & Sahebkar, A. (2023). Rescue effect of curcumin against copper toxicity. Journal of Trace Elements in Medicine and Biology, 127153.

Mohammed, A. A., & Sonawane, K. D. (2022). Destabilizing Alzheimer's Aβ42 protofibrils with oleocanthal: In-silico approach. BIOINFOLET-A Quarterly Journal of Life Sciences, 19(3), 288-295.

Pejaver, V., Urresti, J., Lugo-Martinez, J., Pagel, K.A., Lin, G.N., Nam, H.J., Mort, M. et al. (2020) Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat. Commun. 11, 5918.

Qahtan, O. (2023). Sorting Intolerant from Tolerant and PolyPhen-2 Algorithms: A Variation in Exon 14 of ATP7B Gene among 4 West Iraqi Families with Wilson's Disease. Al-Anbar Medical Journal, 19(2), 98-103.

Takeda, J.-I., Nanatsue, K., Yamagishi, R., Ito, M., Haga, N., Hirata, H., Ogi, T. et al. (2020) InMeRF: prediction of pathogenicity of missense variants by individual modeling for each amino acid substitution. NAR Genom. Bioinform. 2, lqaa038.

Tao, Y.F., Luo, H., Xu, J.B., Cruickshank, A., Zhao, X.R., Teng, F., Hathorn, A. et al. (2021) Extensive variation within the pan-genome of cultivated and wild sorghum. Nat. Plants, 7, 766–773.

Tasmeen, R., Karim, A. B., Banu, L. A., Hossain, E., Rokunuzzaman, M., Majumder, W., ... & Hasan, M. S. (2022). Mutational analysis of exon 8 and exon 14 of ATP7B gene in Bangladeshi children with Wilson disease. Indian Journal of Gastroenterology, 41(5), 456-464.

Verkade, H. J., Bezerra, J. A., Davenport, M., Schreiber, R. A., Mieli-Vergani, G., Hulscher, J. B., ... & Petersen, C. (2016). Biliary atresia and other cholestatic childhood diseases: advances and future challenges. Journal of hepatology, 65(3), 631-642.

Wang, M.J., Zhao, X.M., Takemoto, K., Xu, H.S., Li, Y., Akutsu, T. and Song, J.N.(2012) FunSAV: predicting the functional effect of single amino acid variants using a two-stage random forest model. PLoS One, 7, e43847.

Wang, X.P., Chen, L.M., Liu, W.X., Shen, L.K., Wang, F.L., Zhou, Y., Zhang, Z.D. et al. (2016) AtKC1 and CIPK23 synergistically modulate AKT1-mediated lowpotassium stress responses in Arabidopsis. Plant Physiol. 170, 2264–2277.

Xiong, D., Lee, D., Li, L., Zhao, Q., & Yu, H. (2022). Implications of disease-related mutations at protein–protein interfaces. Current opinion in structural biology, 72, 219-225.

Xu, D., Shao, Q., Zhou, C., Mahmood, A., & Zhang, J. (2023). In silico analysis of nsSNPs of human KRAS gene and protein modeling using bioinformatic tools. ACS omega, 8(14), 13362-13370.

Xu, R., Duan, P.G., Yu, H.Y., Zhou, Z.K., Zhang, B.L., Wang, R.C., Li, J. et al. (2018) Control of grain size and weight by the OsMKKK10-OsMKK4- OsMAPK6 signaling pathway in Rice. Mol. Plant. 11, 860–873.

Yaseen, O. Q., Al-Ani, M. Q., & Majeed, Y. H. (2020). In-Silico prediction of impact on protein function caused by non-synonymous single nucleotide polymorphism in human ATP7B gene associated with Wilson disease. Research Journal of Biotechnology Vol, 15, 3.

Zhao, Q., Feng, Q., Lu, H.Y., Li, Y., Wang, A.H., Tian, Q.L., Zhan, Q.L. et al. (2018) Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. Nat. Genet. 50, 278–284.