

NMR METABOLOME-GENOME WIDE ASSOCIATION STUDY REVEALS GENOMIC REGIONS CONTROLLING GRASS PEA METABOLIC QUALITY: IMPLICATIONS FOR PRECISION BREEDING AND IMPROVED DIETS

Ana Margarida Sampaio¹, Elsa Mecha^{1,2}, Mara Lisa Alves¹, Elsa Brito^{1,2}, Letice Gonçalves¹, Manolis Matzapetakis¹, Iola F. Duarte³, Maria do Rosário Bronze^{1,2,4}, Maria Carlota Vaz Patto^{1*}

¹Instituto de Tecnologia Química e Biológica António Xavier, Oeiras, Portugal

²Instituto de Biologia Experimental e Tecnológica, Oeiras, Portugal

³CICECO – Instituto de Materiais de Aveiro, Departamento de Química, Universidade de Aveiro, Aveiro, Portugal

⁴Faculdade de Farmácia, Universidade de Lisboa, Lisboa, Portugal

*Presenting author: cpatto@itqb.unl.pt

DOI: 10.5281/zenodo.15262277

Growing plant-based eating habits boost legume popularity. Legumes offer rich nutrition and health benefits, having a functional food status. However, nutritional compositional data on some promising legume species is limited, hindering their use. *Lathyrus sativus* (grass pea) is one such understudied species, being an important source of protein and calories in drier areas, with ability to respond to climate challenges (Gonçalves *et al.*, 2022). In this work, untargeted nuclear magnetic resonance (NMR) spectroscopy profiling was used for a broad and unbiased assessment of seed flour metabolite changes of a collection of 200 grass pea accessions, field trialled over three years. NMR assessment of this collection, for which significant genotype-by-sequence information (5,651 SNPs) exist (Sampaio *et al.*, 2021), allowed also to identify the main genomic regions controlling its metabolome variability, through genome wide association study (GWAS). For that, NMR spectra were consolidated into 514 NMR buckets/metabolome features (NMRprocflow, MetaboAnalyst) and their relative concentrations adjusted means calculated using linear mixed models.

From the metabolomic profiles multivariate analysis, a clear year differentiation was observed. 376 NMR features showed genotypic differences, reinforcing their potential for grass pea improvement, and justifying their GWAS. Using a threshold of $-\log_{10}(p)=4$, 435 SNPs associated with 204 NMR features were identified, prioritizing their annotation. Annotation is ongoing by matching spectral information from 1D and 2D spectra to reference compounds in available databases. Amongst identified compounds are amino acids and derivatives, sugars, sugar-alcohols and organic acids.

The combined genetic analysis highlighted multiple loci impacting compounds abundance through a network of interactions, where individual loci may affect more than one compound and vice versa.

The generated knowledge on metabolite variation and respective genetic basis will support precision breeding efforts to answer consumers' nutritional quality concerns, increasing the use of grass pea in diets.

Gonçalves et al (2022) Agronomy 12:6. <http://doi.org/10.3390/agronomy12061324>

Sampaio et al (2021) The Plant Genome 14:e20154. <https://doi.org/10.1002/tpg2.20154>