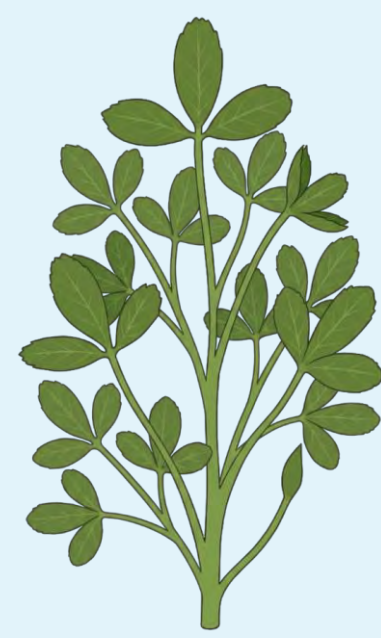




Exploring the alfalfa proteome for optimizing amino acid nutrition



Christina Stonoha-Arther and Kevin Panke-Buisse
U.S. Dairy Forage Research Center, USDA-ARS
Madison, WI 53706

INTRODUCTION

- Legumes are a staple source of dietary protein for humans and animals but are incomplete proteins, *i.e.* they cannot be the sole source of protein
- Alfalfa (*Medicago sativa*) is a forage legume and is the most abundant source of forage protein for dairy cattle. As a perennial forage, alfalfa offers several economic and ecosystem benefits but also presents challenges to overcome
- Low relative quantities of amino acids tied to milk production, e.g. methionine may limit the utility of alfalfa in dairy rations, which, in turn, leads to reduced
- It may be possible to achieve increased methionine levels in legumes by cisgenic overexpression of native methionine-rich genes or motifs
- We have developed an R package, residu, that facilitates the characterization and organization of peptide sequences based on amino acid content
- Exploration of the alfalfa proteome with residu and other proteomics tools to better characterize the amino acid profile of alfalfa revealed potential for further investigation
- Particular emphasis was placed on methionine- and cysteine content of alfalfa proteins and comparison to corn as the predominant forage crop

METHODS

- The latest available versions of the alfalfa (Shen et al. 2020) and maize B73 (v5.0) proteomes were used for all analyses
- The current alfalfa proteome did not capture an ORF corresponding to the RuBisCo large subunit so the *Medicago truncatula* (v4.0) sequence was used as the reference sequence to identify it in the alfalfa genome
- UniProt, BLAST, and the R package, residu, were used for all analyses

RESULTS

Table 1. Summary of amino acid groups and key individual amino acids in the alfalfa proteome.

amino acid group (or specific amino acid)	percent		count	
	mean	median	mean	median
sulfur amino acid group	4.76	4.49	16.77	13
methionine	2.79	2.58	9.60	7
cysteine	1.97	1.75	7.17	5
positive amino acid group	14.32	14.17	52.61	41
lysine	6.94	6.60	25.48	19
negative amino acid group	11.90	11.76	44.80	34
hydrophobic amino acid group	49.08	49.28	181.85	143
polar amino acid group	52.89	52.75	197.64	154
aromatic amino acid group	11.35	11.33	41.77	33
aliphatic amino acid group	30.80	30.92	114.34	89

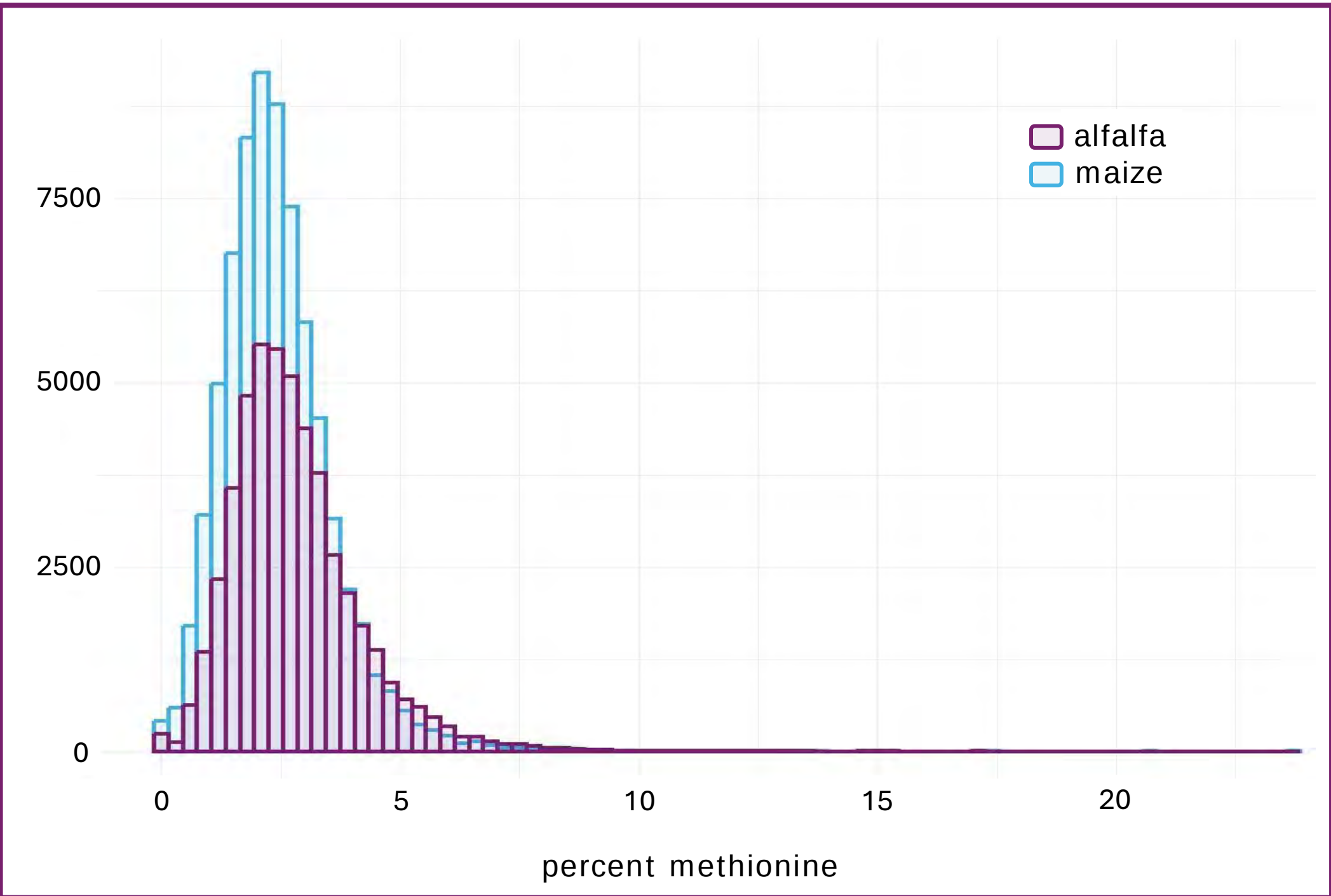


Figure 1. Histogram of methionine content by percent of the alfalfa (purple) and maize (blue) proteome generated with residu and ggplot2.

RESULTS

Table 2. Percent cysteine and methionine of the top ten protein sequences by sulfur-containing amino acids in alfalfa.

rank	sequence name	Percent sulfur group amino acids	percent cysteine	percent methionine
1	MsG0680033529.01	28.81	27.97	0.85
2	MsG0180004392.01	25.37	16.42	8.96
3	MsG0680034934.01	24.32	7.21	17.12
4	MsG0780038735.01	23.08	17.95	5.13
5	MsG0280010194.01	22.64	20.75	1.89
6	MsG0880042211.01	21.92	20.55	1.37
7	MsG0280010708.01	21.82	14.55	7.27
8	MsG0180000301.01	21.59	13.64	7.95
9	MsG0580027919.01	21.07	8.81	12.26
10	MsG0880044237.01	20.29	13.04	7.25

Table 3. Percent methionine and methionine density of the top ten protein sequences by methionine content in alfalfa.

rank	sequence name	percent methionine	Maximum methionine per 100 amino acids
1	MsG0680034934.01	17.12	19
2	MsG0880046964.01	15.15	15.51
3	MsG0080048523.01	15.00	15
4	MsG0780040155.01	14.67	20
5	MsG0280009057.01	14.61	23
6	MsG0680035861.01	13.43	13.43
7	MsG0280010292.01	13.16	13
8	MsG0480023438.01	13.04	13.04
9	MsG0880046822.01	13.04	13.04
10	MsG0580025778.01	12.96	12.96

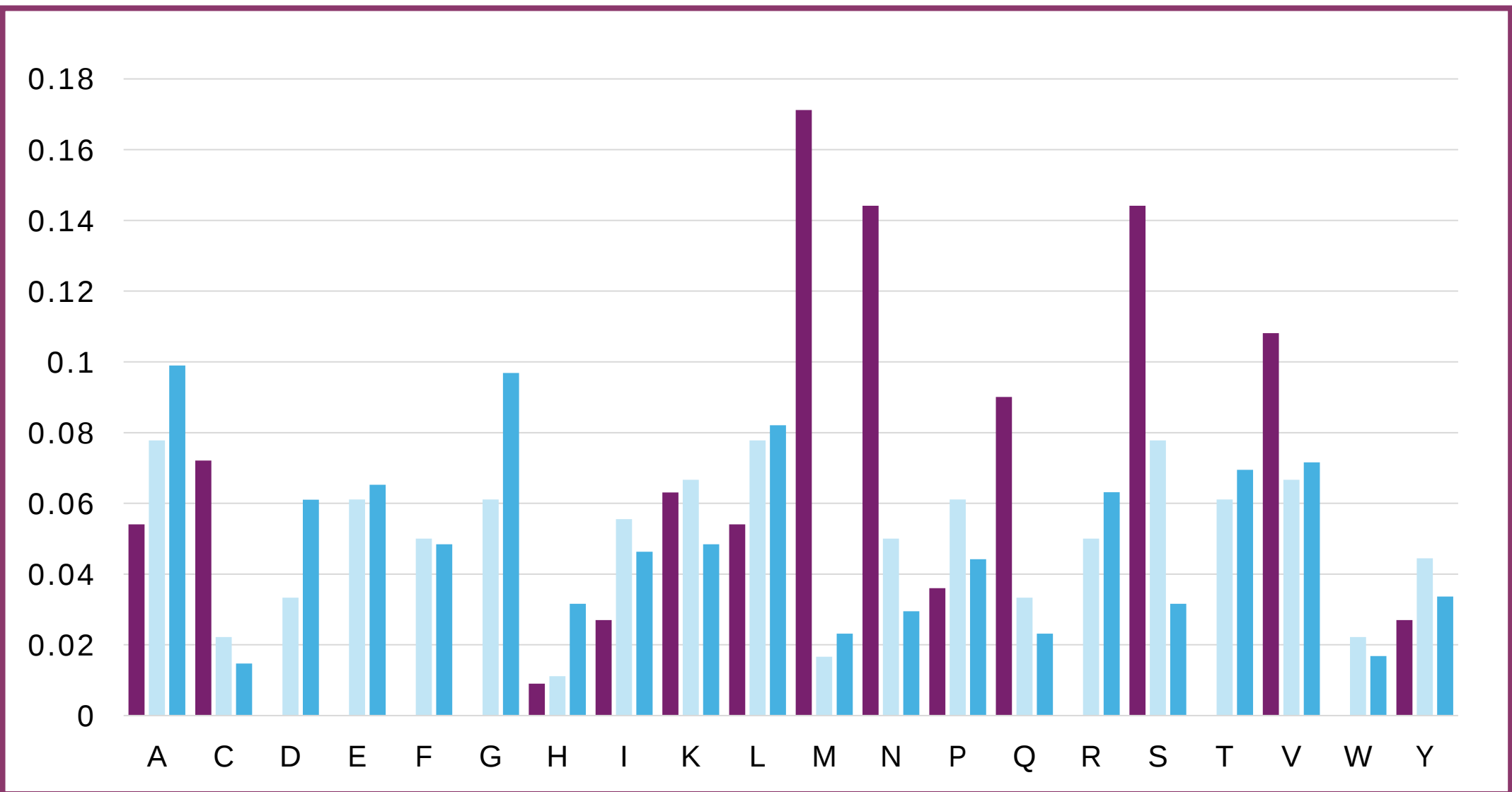


Figure 2. Amino acid frequency plot, generated with residu, of the alfalfa proteome (purple) beside the small (light blue) and large (dark blue) subunits of RuBisCo, the predominant dietary protein in alfalfa.

- There is a wide range of methionine and sulfur-containing amino acid content within the alfalfa proteome
- The corn proteome is larger than the alfalfa proteome, but no clear differences in methionine content across the proteome are apparent
- Sulfur group amino acid content tracks more closely with cysteine content with only the highest methionine content sequence also appearing in the top ten sulfur group
- There are several cisgenic candidates for potential overexpression or modification within the alfalfa proteome to increase methionine content
 - The top ten highest methionine sequences range from 12% - 17% and the highest density is 23%
- While the overall alfalfa proteome is rich in methionine, the most abundant protein for ruminant dietary intake, RuBisCo, is quite low

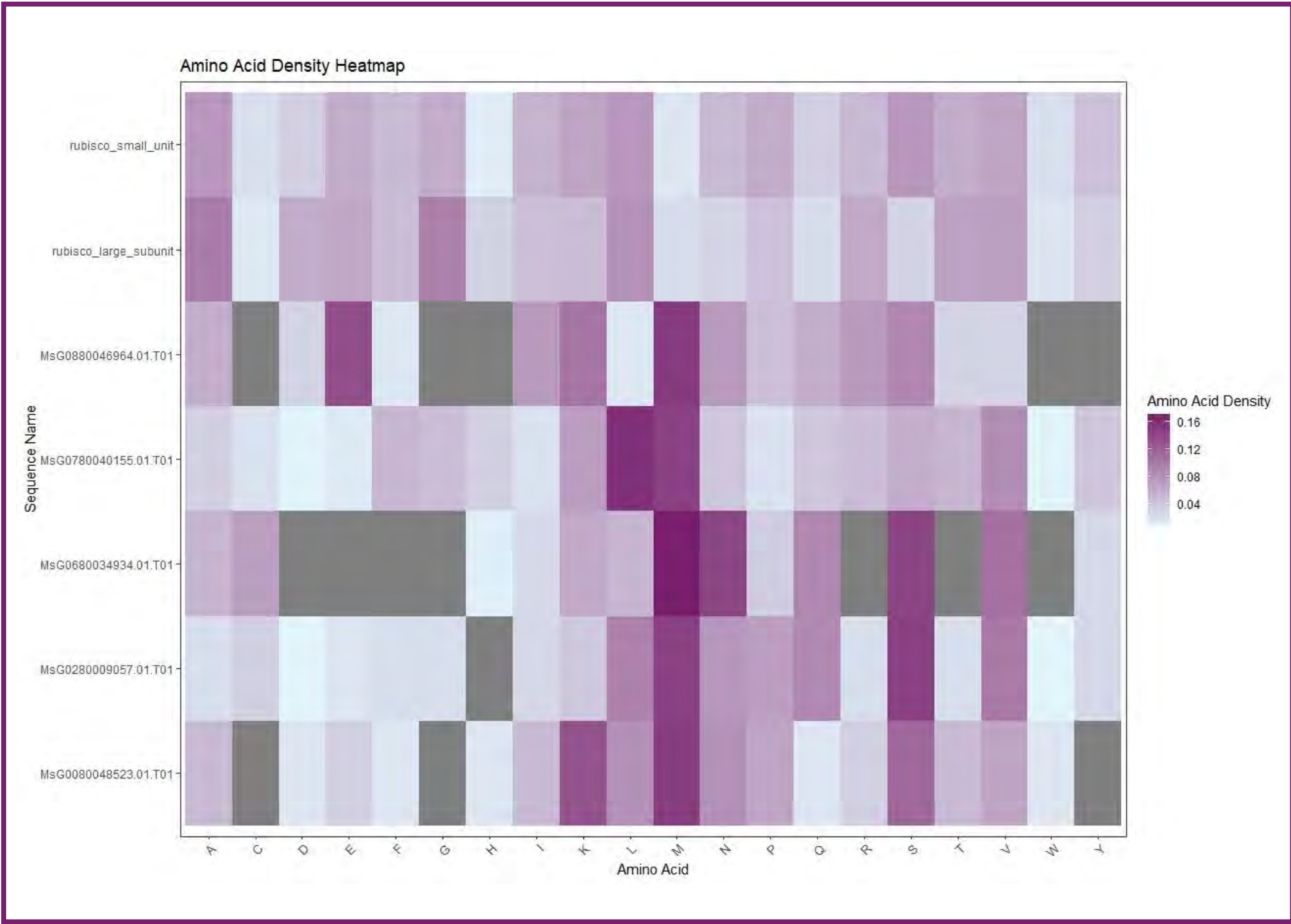


Figure 3. Amino acid density heatmap, generated with residu, of the top five highest percent methionine protein sequences from alfalfa and the small and large subunits of RuBisCo. Shades of purple indicate amino acid frequency/density. Grey boxes indicate the amino acid is not present in the sequence.

DISCUSSION

- Improvement of alfalfa amino acid nutrition is one strategy for increasing alfalfa acreage and mitigating protein degradability challenges
- Preliminary evaluation indicates strong potential for cisgenic nutritional augmentation of alfalfa protein
- In particular, these results highlight that nutritional dietary amino acid content is not driven by the overall proteome, but by the abundance of individual proteins
- Next steps include:
 - Expansion of genomic and proteomic resources for alfalfa to improve prediction, annotation, and narrow the field of genetic modification candidates
 - Model the impacts of nutritional optimization at farm and regional scales
 - Profile high-density motifs for target amino acids



DOWNLOAD

residu is available and maintained on GitHub and Zenodo by the authors

Feedback, suggestions, or questions can be emailed to: christina.arther@usda.gov

GitHub:



Zenodo:



Every download helps

REFERENCES

- Stonoha-Arther C, Panke-Buisse K. (2025) residu: An R package for exploring amino acid residue composition of peptide and protein sequences. *Frontiers in Horticulture* 4:1686134. <https://doi.org/10.3389/fhort.2025.1686134>.
- Shen C, Du H, Chen Z et al. (2020) The chromosome-level genome sequence of the autotetraploid alfalfa and resequencing of core germplasms provide genomic resources for Alfalfa research. *Molecular Plant* 13(9): 1250-1261.
- Margaret R Woodhouse, Ethalinda K Cannon, John L Portwood 2nd, Jack M Gardiner, Rita K Hayford, Olivia Haley, Carson M Andorf. (2025) Tools and Resources at the Maize Genetics and Genomics Database (MaizeGDB). *Cold Spring Harb Protoc.* 2025(1). 10.1101/pdb.over108430

