

The Unique Protein Composition of Honey Revealed by Comprehensive Proteomic Analysis: Allergens, Venom-like Proteins, Antibacterial Properties, Royal Jelly Proteins, Serine Proteases, and Their Inhibitors

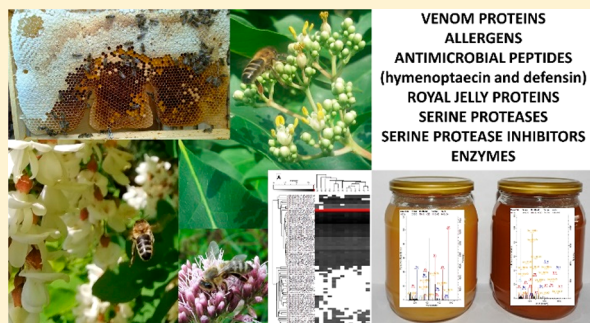
Tomas Erban,^{*,†,‡} Elena Shcherbachenko,[†] Pavel Talacko,[‡] and Karel Harant[‡]

[†]Proteomics and Metabolomics Laboratory, Crop Research Institute, Drnovska 507/73, Prague 6-Ruzyne, CZ-16106, Czechia

[‡]Proteomics Core Facility, Faculty of Science, Charles University, BIOCEV, Prumyslova 595, Vestec, CZ-25242, Czechia

S Supporting Information

ABSTRACT: Honey is a unique natural product produced by European honeybees. Due to its high economic value, honey is considered to be well characterized chemically, and it is often discovered to be an adulterated commodity. However, this study shows that our knowledge of honey protein composition, which is of high medical and pharmaceutical importance, is incomplete. In this in-depth proteomic study of 13 honeys, we identified a number of proteins that are important for an understanding of honey properties and merit additional pharmaceutical research. Our major result is an expanded understanding of the proteins underlying honey's antimicrobial properties, such as hymenoptaecin and defensin-1, glucose dehydrogenase isoforms, venom allergens and other venom-like proteins, serine proteases and serine protease inhibitors, and a series of royal jelly proteins. In addition, we performed quantitative comparisons of all of the proteins previously known or newly identified. The honey proteins, determined using label-free nLC-MS/MS in which the same protein quantity was analyzed in one series, were found in relatively similar proportions, although eucalyptus honey differed most widely from the remaining honeys. Overall, the proteome analysis indicated that honeybees supply proteins to honey in a relatively stable ratio within each proteome, but total protein quantity can differ by approximately an order of magnitude in different honeys.



Honey is one of the unique products provided by the activity of European honeybees, *Apis mellifera* Linnaeus, 1758.¹ Furthermore, beekeeping is a part of many cultural heritages, and in some cultures, honey has religious significance.² Ancient and recent civilizations have used honey as an important food source, but honey also has high and diverse medical potential, including a specific pharmaceutical application as an antimicrobial agent³ targeting multidrug-resistant microbes.^{4,5} To limit antibiotic use, the recently released draft guidance of the National Institute for Health and Care Excellence (NICE) indicates that honey can serve as an official self-care remedy recommended by doctors for topical treatment of upper respiratory tract diseases.⁶ Additional potential medical applications of honey include treating burns and wounds,^{7–9} dental care,¹⁰ and use as a natural antitumor agent.^{11,12} Furthermore, honey is a highly valued commodity, but its price can vary considerably across countries.¹³ Honey is considered to be well characterized chemically, partially because it is often adulterated;¹⁴ however, its protein composition has not been well characterized, and knowledge of its protein components can open new avenues for medical/pharmaceutical applications of honey.

Relatively few studies exist regarding protein analysis of honey because the proteins occur in low quantities of approximately 0.1–0.5%, making protein extraction difficult,¹⁵ and because protein quantity is influenced by the origin of the honeybees.¹⁶ Despite the minute levels of proteins in honey, knowledge of honey proteins is important from several perspectives, particularly honey's high potential in pharmacology^{5–12,17} and the use of proteins for honey authentication,¹⁴ a consistently challenging but solvable problem. Honey is known to exhibit acid phosphatase activities and contains the following enzymes: α , β -amylase or diastase; α -glucosidase, also known as invertase; glucose oxidase.¹⁸ These major enzymes are primarily produced by honeybee glands.^{19–21} Furthermore, catalase activity occurs in honey, but the origin of this enzyme activity is unclear.²² Analyses using SDS-PAGE bands or spots have shown that protein patterns can be used to distinguish honey proteins.^{23,24} Won et al.²⁵ used SDS-PAGE slices to identify subunits of major royal jelly protein 1 (MRJP1) as the prevalent protein in honey.²⁵ Using 2D-E-MALDI-MS, Rossano et al.²⁶ identified MRJPs 1–5, alpha-

Received: November 16, 2018

Table 1. Results of the nLC-MS/MS Analysis, Which Includes 71 Protein Hits Resulting from Matrix Reduction to at Least 3 Positive Results

	Intensity	MS/MS Count	Honey													Honey	Venom	Royal jelly	Exact allergen	Allergen analog	Signal peptide
			H01	H02	H03	H04	H05	H06	H07	H08	H09	H10	H11	H12	H13						
1E+09	74	25.1	26.2	25.6	25.7	25.1	23.2	25.6	25.5	25.9	25.3	25.8	24.5	25.4	Def1	NP_001011616.2	defensin-1	X	X	X	
6E+07	26	NaN	NaN	23.3	NaN	NaN	NaN	NaN	22.0	NaN	NaN	NaN	NaN	23.3	LOC406142	NP_001011615.1	hymenoptaecin	X	X	X	
5E+10	1808	30.0	30.7	30.9	29.9	29.3	30.6	30.8	29.7	30.3	29.4	30.3	29.5	30.4	LOC406081	BAA86908.1	glucose oxidase (notatin)	X	X	X	
4E+08	44	24.2	25.6	NaN	NaN	23.8	24.1	NaN	24.4	NaN	22.7	25.2	NaN	24.6	GMCOX7	XP_006565048.1	glucose dehydrogenase [FAD, quinone]			X	
3E+09	556	26.1	26.2	26.2	26.6	25.8	26.6	26.4	26.0	27.1	25.8	26.2	25.5	26.1	Gld2	XP_016767722.1	glucose dehydrogenase [FAD, quinone]	+	X		
1E+10	1541	28.2	28.2	28.8	28.4	27.7	28.6	28.9	28.1	28.8	27.4	28.5	28.1	28.7	GMCOX14	XP_003251149.1	glucose dehydrogenase [FAD, quinone]	+	X	X	
1E+08	106	21.8	21.8	21.9	22.4	21.9	22.4	22.2	22.2	21.7	21.9	22.1	21.9	21.8	GMCOX12	XP_394210.5	glucose dehydrogenase [FAD, quinone]			X	
2E+08	136	22.4	23.3	23.4	23.1	22.1	22.8	22.9	22.6	22.2	21.6	23.1	22.8	23.4	LOC100576326	XP_016769186.2	phenoloxidase-activating factor 2			X	
1E+09	270	24.4	25.0	24.5	24.8	24.4	25.1	25.2	25.0	25.2	24.3	24.8	24.9	25.5	LOC410149	XP_393632.2	carboxypeptidase Q	+	X	X	
4E+09	525	26.9	27.1	27.5	27.0	26.7	27.1	27.3	26.7	26.7	26.5	27.3	26.6	27.0	SP3	XP_006560620.1	CLIP / venom serine protease Bi-VSP	+	X	X	
3E+09	731	24.8	26.1	26.4	26.3	25.8	26.2	26.6	26.0	26.6	26.1	26.5	27.0	25.6	SP28	XP_392669.1	CUB / venom serine protease 34 (Api m 7 analog)	+	X	X	
5E+07	52	21.4	NaN	21.2	21.5	20.5	21.2	21.3	NaN	NaN	NaN	21.0	NaN	20.9	LOC406146	AAA27730.1	hyaluronidase (Api m 2 allergen)	X	X	X	
6E+07	37	22.3	NaN	NaN	NaN	NaN	22.1	21.9	NaN	NaN	22.1	NaN	NaN	21.7	LOC725215	XP_006569959.1	venom acid phosphatase Acph-1-like (Api m 3 analog)	X	X	X	
4E+08	176	23.3	23.5	24.2	23.6	23.1	22.9	23.1	22.9	22.6	22.9	23.2	22.7	24.2	LOC726737	XP_006569974.1	venom acid phosphatase Acph-1-like (Api m 3 analog)	X	X	X	
1E+09	160	24.9	25.1	25.5	24.4	24.2	25.1	25.1	24.6	25.0	23.5	24.5	23.9	25.5	LOC503505	ABF21078.1	icarapin (Api m 10 allergen)	X	X	X	
6E+08	193	23.6	24.1	24.4	24.1	23.9	24.3	24.3	23.8	24.3	23.2	24.1	23.7	24.2	LOC413324	XP_016769016.1	chitinase-like protein Idgf4	?	X	X	
2E+09	249	25.6	25.5	26.2	26.0	25.2	25.8	25.8	24.7	25.9	25.5	25.9	25.6	25.4	Mrjp9	AAV21180.1	MRJP9 (Api m 11 allergen)	X	X	X	
9E+08	372	24.5	24.5	25.4	25.2	24.6	25.1	24.8	24.5	24.5	24.3	24.6	24.5	24.5	Tsf1	NP_001011572.1	transferrin I	X	X	X	
9E+07	38	NaN	NaN	NaN	NaN	NaN	21.3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	RfaBp	XP_026298285.1	apolipoporphins	X		X	
1E+08	62	22.4	NaN	22.2	22.3	NaN	NaN	22.3	NaN	22.3	21.8	NaN	NaN	22.7	LOC410620	NP_001153544.1	Hsc70-3 / Hsp70Ab	X		X	
3E+07	16	NaN	NaN	NaN	23.7	NaN	NaN	NaN	NaN	23.5	NaN	23.3	NaN	NaN	LOC102655073	XP_006570863.1	histone H4	X			
2E+08	51	NaN	NaN	NaN	NaN	23.0	NaN	NaN	22.8	NaN	NaN	NaN	22.5	22.8	Rpl40	XP_003251367.1	ubiquitin-60S ribosomal protein L40	X			
1E+08	40	NaN	NaN	25.0	22.7	22.1	NaN	NaN	NaN	21.9	NaN	NaN	NaN	22.2	Ef-1a-f1	CAA37066.1	elongation factor 1-α	X			
1E+09	286	24.3	25.3	25.4	24.9	24.7	24.9	24.7	25.3	24.5	25.4	24.7	24.6	25.0	LOC725404	XP_026296530.1	actin				
1E+09	97	23.6	NaN	25.0	26.4	26.1	25.1	NaN	26.0	NaN	23.2	25.9	25.8	25.5	LOC406114	BAA86909.1	α-amylase	X	X	X	
3E+10	1323	29.5	29.8	29.7	29.7	29.0	29.6	29.9	28.8	29.7	29.7	29.9	29.5	30.0	LOC406114	AAM20738.1	α-amylase			X	
1E+10	372	27.4	27.8	28.3	26.6	26.8	27.8	28.1	27.4	25.5	28.4	27.8	27.9	28.8	Hbg3	Q17058.1	alpha-glucosidase III	X	X	X	
2E+11	3121	32.1	32.5	33.0	32.1	31.4	32.2	32.6	31.7	32.7	32.5	32.5	32.2	32.9	Hbg3	BAE86928.1	alpha-glucosidase III	X	X	X	
2E+12	8162	35.5	35.4	35.3	35.4	34.9	35.8	35.7	35.3	35.6	34.8	35.5	34.9	35.6	Mrjp1	NP_001011579.1	MRJP1	X	X	X	
3E+11	3489	33.3	33.0	33.2	33.5	32.9	33.8	33.6	33.1	33.4	32.6	33.4	33.1	33.7	Mrjp2	AQM49880.1	MRJP2	X	X	X	
3E+10	102	NaN	29.8	29.6	30.4	29.8	28.4	30.8	30.2	30.2	29.1	30.4	29.4	30.3	Mrjp2	NP_001011580.1	MRJP2	X	X	X	
3E+11	4374	32.9	32.2	32.4	33.1	32.2	33.3	33.0	32.6	32.0	31.8	32.8	32.4	33.2	Mrjp3	ADC55524.1	MRJP3	X	X	X	
6E+09	210	27.3	28.0	27.3	27.4	27.5	27.8	27.7	27.6	NaN	27.6	27.7	27.4	28.2	Mrjp4	NP_001011610.1	MRJP4	X	X	X	
2E+10	1095	29.0	29.1	29.5	29.6	28.9	29.6	29.4	29.0	29.4	28.6	29.6	29.3	29.7	Mrjp4	XP_026299317.1	MRJP4	X	X	X	
3E+09	73	26.4	26.7	NaN	26.2	25.6	27.1	26.8	NaN	NaN	25.8	26.3	25.4	26.7	Mrjp5	NP_001011599.1	MRJP5	X	X	X	
1E+11	3071	31.5	31.3	31.6	31.7	30.9	31.6	31.9	30.8	31.2	31.0	31.5	31.2	31.6	Mrjp5	ADB96941.1	MRJP5	X	X	X	
5E+09	479	26.5	26.3	27.2	26.7	25.5	26.6	26.1	25.5	28.4	25.5	27.0	26.6	25.8	Mrjp5	XP_026299316.1	MRJP5	X	X	X	
2E+10	995	28.8	29.6	29.9	29.1	28.6	29.2	29.5	28.5	27.7	29.1	29.1	29.0	29.5	Mrjp6	NP_001011622.1	MRJP6	X	X	X	
5E+10	1836	30.8	30.4	30.5	30.7	30.0	30.9	31.0	30.4	30.1	29.9	30.6	30.2	30.5	Mrjp7	NP_001014429.1	MRJP7	X	X	X	
9E+10	710	31.5	31.9	31.7	30.7	29.4	31.1	31.7	31.1	31.5	30.0	31.0	29.5	31.9	LOC408608	XP_397512.1	uncharacterized protein LOC408608	X	X	X	
3E+08	77	23.1	23.8	24.1	NaN	NaN	23.1	23.6	22.9	24.9	NaN	NaN	NaN	23.1	A4	NP_001107670.1	apolipoporphin-III-like	X		X	
3E+07	36	20.9	NaN	21.0	NaN	NaN	NaN	20.2	NaN	NaN	NaN	NaN	NaN	21.1	LOC726323	XP_026301350.1	uncharacterized protein LOC726323			X	
5E+07	29	NaN	NaN	NaN	NaN	22.5	21.8	NaN	NaN	23.0	NaN	NaN	NaN	NaN	Obp14	NP_001035313.1	Obp 14	X	X	X	
3E+07	33	21.2	NaN	NaN	NaN	NaN	21.1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	LOC727028	XP_001122741.2	unch. protein LOC727028 / takeout-like			X	
2E+08	100	22.4	22.8	22.5	NaN	NaN	22.4	22.4	NaN	22.5	21.9	21.7	NaN	22.8	Fer2LCH	XP_624076.1	ferritin heavy polypeptide-like 17	X	X	X	
2E+08	50	23.6	NaN	22.0	NaN	NaN	23.6	NaN	NaN	21.6	NaN	NaN	NaN	23.8	Fer1HCH	XP_026300991.1	ferritin subunit (heavy chain)	X		X	
2E+09	266	25.9	25.9	26.7	26.2	25.2	25.8	26.1	25.9	24.9	25.4	26.0	25.7	25.9	LOC552811	XP_625189.3	laccase-5 (? polyphenol oxidase)			X	
9E+09	679	27.9	28.5	28.2	27.7	26.9	27.9	28.4	27.4	28.6	26.8	27.8	26.6	28.8	LOC552229	XP_016770320.2	esterase B1	+		X	
5E+08	70	24.0	NaN	23.7	23.5	NaN	23.7	23.7	23.6	25.1	23.4	24.0	NaN	23.0	LOC551268	XP_026299639.1	pancreatic triacylglycerol lipase			X	
2E+10	506	28.9	29.7	29.8	28.3	27.5	29.3	29.3	28.4	29.0	28.2	28.6	27.5	29.0	LOC727193	XP_026299638.1	lipase member H-A	+	X	X	
4E+08	141	24.0	24.5	24.4	NaN	NaN	23.7	23.8	23.7	24.5	NaN	23.3	NaN	23.0	LOC551268	XP_623663.1	pancreatic triacylglycerol lipase	+		X	
8E+09	685	27.9	28.1	27.9	28.0	27.5	28.1	28.5	27.7	28.0	27.5	28.0	27.9	28.4	LOC413627	XP_026298083.1	unch. protein LOC413627 (? regucalcin Blastp)			X	
3E+09	390	26.4	26.5	26.3	26.2	25.6	26.7	26.6	26.0	27.3	25.2	26.3	25.5	26.0	LOC409708	XP_006570612.1	glucosylceramidase-like	X	X	X	
2E+09	246	25.7	25.7	25.9	25.9	25.2	25.1	25.5	25.3	26.2	25.2	25.6	25.5	25.7	LOC409709	XP_393208.1	putative glucosylceramidase 4	X	X	X	
3E+09	207	26.0	26.8	26.2	25.1	24.4	26.1	26.3	25.9	27.6	25.2	25.6	24.5	26.8	LOC102655185	XP_026299771.1	uncharacterized protein LOC102655185			X	
1E+09	571	25.2	25.2	25.6	25.3	24.0	25.5	25.7	24.3	25.0	23.1	24.9	24.1	25.3	LOC724933	XP_016768886.2	xanthine dehydrogenase			X	
8E+07	89	21.6	21.6	21.4	21.9	NaN	22.0	22.0	21.7	NaN	NaN	22.0	21.8	21.5	LOC552249	XP_026296064.1	lysosomal alpha-mannosidase			X	
2E+08	97	22.2	NaN	23.0	22.7	22.2	22.4	22.5	22.4	22.1	21.7	22.2	22.3	22.8							

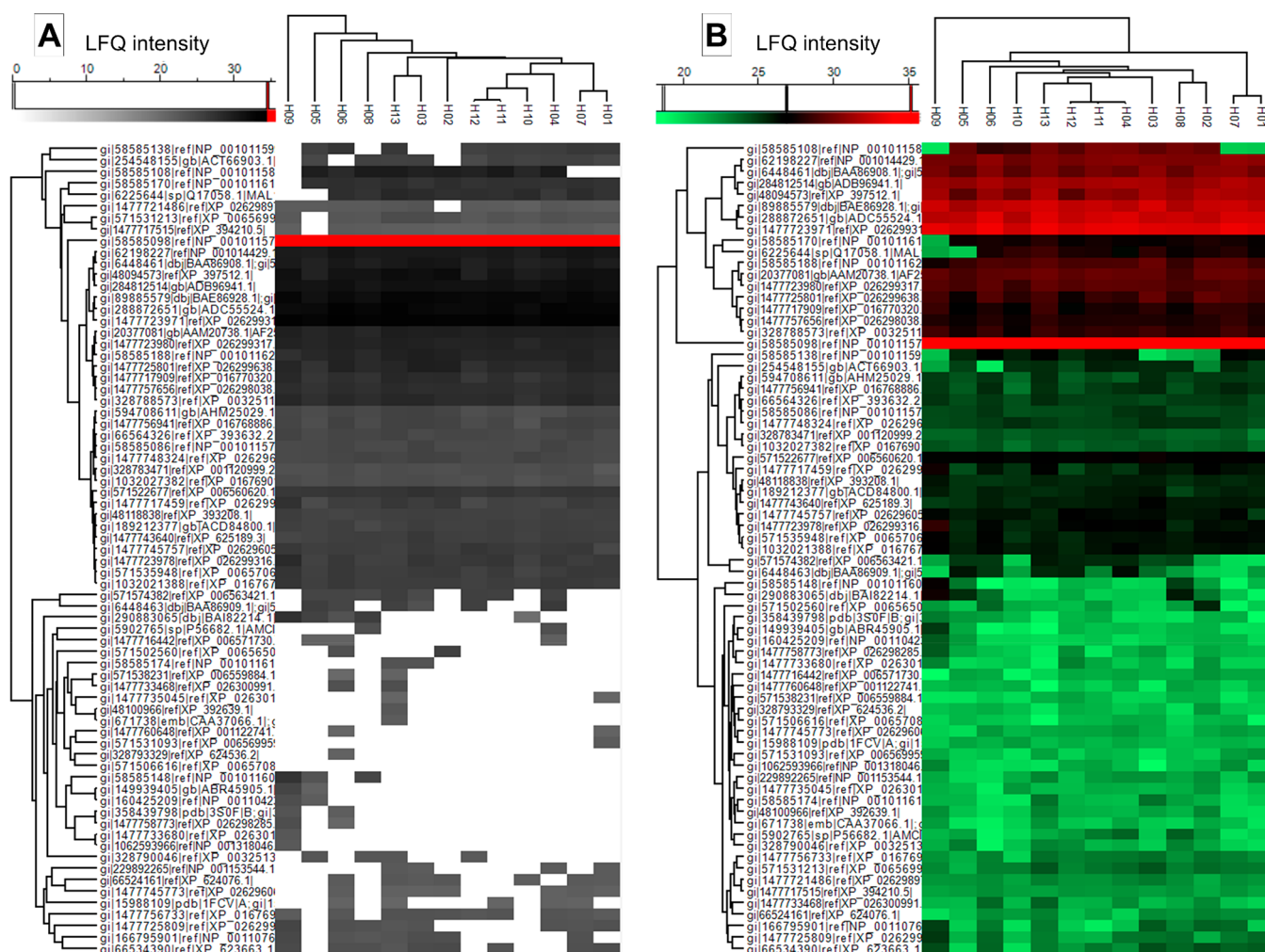


Figure 1. Two heatmaps that illustrate proteome differences in proteins of the 13 honeys. (A) A heatmap resulted after the missing values were replaced by the constant “0”. The heatmap clearly shows the missing values (in white) and MRJP1 protein (in red; gil58585098), which was of highest abundance. (B) A heatmap was generated after the missing values were replaced according to a normal distribution. The heatmap well represents data comparison for proteins that were detected among the samples, but the qualitative result is less visible, and again, MRJP1 is most intense and clearly visible.

al.²⁸ identified MRJPs 1, 2, 5, and 7. In recent studies, Zhang et al.²⁹ and Borutinskaite et al.³⁰ reported identification of MRJPs 1–9, α -glucosidase, α -amylase, glucose oxidase, and uncharacterized protein LOC408608. The latter study, which analyzed buckwheat honey, identified additional proteins such as transferrin 1, glucosylceramidase-like protein, and putative glucosylceramidase 4,³⁰ and the authors reported additional proteins that were denoted as uncharacterized (see Table 1 in this study). Interestingly, Azevedo et al.³¹ aimed to discriminate among honeys using 2D-E proteomics.

Here, we describe qualitative and quantitative differences among different honeys and, importantly, identify a number of previously unknown proteins in honey. This study greatly increases our knowledge of the protein composition of honey and provides the basis for understanding the unique and multiple medicinal properties of honey. Additionally, identification of allergens and allergen analogues is useful for allergy treatment and research. Importantly, these results should also be applicable to honey authentication. The results were achieved by analyzing 13 different authentic honeys using the label-free quantification (LFQ) proteomic approach by employing nanoliquid chromatography (nLC) coupled with

the state-of-the-art Orbitrap Fusion Tribrid mass spectrometer. Because the major components of honey are sugars—mainly glucose and fructose¹⁴—we performed complementary sugar analysis in the investigated honeys using HPLC coupled with an evaporative light-scattering detector (ELSD). The total protein content was also calculated for each honey.

RESULTS AND DISCUSSION

The gel filtration technique using Sephadex G-25 medium is a good cleanup strategy for proteomic analysis of honey. The primary data evaluation from the 39 consecutive nLC-MS/MS runs enabled the identification of 119 protein hits assigned taxonomically to honeybee, and after filtering for at least three positive results in total, this data set was reduced to 71 hits (Table 1, Table S1 Supporting Information); however, the list includes some isoforms. As indicated in Table 1, most of the proteins are likely secreted as a signal peptide, indicated using a combination of bioinformatic tools.^{32–34} Proteins for which we failed to identify a signal peptide can be hypothesized to enter honeys through regurgitation of entire cells. In some cases, it is possible that the current tools could not identify the site, but the list of identified proteins suggests that the former

Table 2. Results of the Sugar Analysis and Total Proteins in Honey^a

honey sample		protein content in honey			fructose and glucose contents in honey				
name	no.	[mg/g]	±SD	%	fru [mg/g]	±SD	glu [mg/g]	±SD	fru/glu
linden	H01	0.3	0.1	0.03	392	32	287	23	1.42
buckwheat	H02	1.4	0.2	0.14	385	8	316	4	1.25
sunflower	H03	0.4	0.1	0.04	431	16	345	8	1.28
acacia	H04	0.1	0.0	0.01	448	11	298	8	1.55
acacia	H05	0.1	0.0	0.01	485	11	336	2	1.48
wild bees	H06	0.6	0.1	0.06	384	9	336	4	1.17
honey of the year	H07	0.5	0.1	0.05	478	9	381	8	1.28
linden	H08	0.2	0.0	0.02	402	16	342	6	1.20
eucalyptus	H09	0.9	0.0	0.09	383	13	220	5	1.84
fruit trees-apple	H10	0.1	0.0	0.01	394	4	344	6	1.17
linden	H11	0.3	0.1	0.03	410	22	342	5	1.22
floral	H12	0.1	0.0	0.01	397	15	384	7	1.05
black-forest	H13	0.3	0.0	0.03	353	10	289	8	1.26

^aThe proteins were determined using the Bradford method with bovine serum albumin as standard after sample cleaning by PD MidiTrap G-25 columns. Sugars were determined using the HPLC-ELSD technique.

explanation is more likely in most cases. Finally, some proteins can be excreted via an unconventional protein secretion route.

The heatmaps in Figure 1 illustrate the LFQ results and show the relationships among the 13 honey proteomes. We justify the data evaluation using the two algorithms in Perseus;³⁵ one heatmap (Figure 1A) better demonstrates the qualitative data position, while the second heatmap (Figure 1B) is common in comparisons of expression data, but the qualitative outcome of the data is hidden. The heatmap visualization indicated a similar distribution of the honeybee proteins among the analyzed honeys when the same protein concentration is analyzed; however, some differences were observed in the cluster analysis. The cluster analysis in both heatmaps clearly shows that eucalyptus honey (H09) is the sample for which the proteome differs most widely from those of the remaining 12 honeys, and the difference might be influenced by the Spanish origin of the honeybees¹⁶ because only the eucalyptus honey was not a product of Czechia, where the *A. mellifera carnica* subspecies prevails. Additionally, forest honey (H13), which is prepared by bees from the excretions of aphids, contained similar honeybee proteins to those in the remaining 12 honeys from the plant excretions.

For the data interpretation, it was very important to realize that the samples analyzed using nLC-MS/MS had the same protein content. Thus, the main output that we obtained from these data, other than the qualitative results (detected and missing proteins), is the variation in protein abundance among the honeys when they are hypothetically dissolved to the same protein concentration. Importantly, the majority of the proteins did not show high quantitative variations among the samples (vertical view), but consistent with previous reports, we found relatively substantial variations in the total protein contents among the honeys. The total protein quantities determined in the honeys (Table 2) were mostly in the range of tenths of mg per g of honey,¹⁵ but the protein levels in two honeys were approximately 1 mg per g of honey: buckwheat (H02) and eucalyptus (H09) honey. A relatively high protein content in eucalyptus honeys (0.91–1.24 mg per g of honey) was also found in a previous study by Rossano et al.²⁶ The protein mixtures in the label-free nLC-MS/MS analysis were analyzed at the same concentration, and the LFQ results obtained indicate no dramatic quantitative differences among samples. Thus, although the bees supply honey with a relatively

stable protein mixture with respect to the proportion of proteins in the proteome, the total protein quantity can differ by approximately one order of magnitude; herein, this quantity is the key factor in producing differences in the particular proteins/enzymes in the honey. Note that the gel filtration technique removed salts, sugars, and other low-molecular-weight compounds. Furthermore, the nondestructive cleanup procedure excluded pollen particles. Thus, the determined total protein content may be lower than that in some other studies.

Further individual examination of the 71 *A. mellifera* protein results revealed the presence and LFQ of important proteins and a number of previously unreported proteins in honey. Notably, we consider the additional 48 protein hits to be trace results (see Table S1, Supporting Information). Most notable among the newly reported honey proteins in this study were hymenoptaecin, venom-related and venom-like proteins, royal jelly related proteins, proven allergens, serine proteases, inhibitors of serine proteases, and isoforms of glucose dehydrogenase (Gdh) [FAD, quinone]. Despite the wide array of proteins identified compared to Rossano et al.²⁶ we were not able to detect profilin, superoxide dismutase, and apisimin. Furthermore, despite detailed analyses performed together with individual investigations of records in databases, we were also unable to match our results for some proteins to those reported in other studies.^{28,30} Moreover, our manual verification of the *A. mellifera*-related proteins reported in Borutinskaite et al.³⁰ revealed that more proteins can be identified in buckwheat honey, but the proteins were denoted as uncharacterized (see Table 1; proteins denoted with + in the column ^aHoney).

The protein array of honeys identified in this study may help explain the multiple curative properties of honey. Furthermore, the presence of venom allergens in honey can contribute to understanding honey allergies.^{36,37} On the other hand, the positive effects of honeybee proteins in honey on allergy prevention are also of importance. Overall, compared to previous studies, this study greatly increases our knowledge of the protein composition of honey. The protein groups are further discussed in separate sections, and a list of proteins with details of the analysis is provided in Table 1. Further details, including details such as conserved domains (CCDs)^{38,39} and all protein IDs, are provided in Table S1,

Supporting Information. To easily compare our data with those of previous studies, we included the database protein searches' gi numbers, a method that is currently not commonly used. This step particularly facilitated the orientation in the previous studies containing the obsolete protein IDs. The complementary results, such as the total protein contents and sugars in each honey, are also presented in Table 2, and the sample chromatograms of each honey (H01–H13) are provided in Figure S1.

Antimicrobial Peptides. An important characteristic of honey is its antibacterial properties, which are increasing further in importance due to the developing resistance of bacteria to antibiotics.^{4,5,40} Therefore, one of the important results of this study is the detection of defensin-1 and hymenoptaecin. While defensin-1 was detected in honey using antibodies⁴⁰ and proteomics,²⁷ this study is the first to report the detection of hymenoptaecin. Additionally, jelleins 1, 2, and 4 in MRJP1 have been shown to contribute to bacteriolysis in honey.⁴¹ We found a constant protein abundance of MRJP1 in the honey proteomes, and we discuss this protein in the section devoted to royal jelly. The fact that MRJP1 had the highest abundance in all honeys suggests the high potential of jelleins 1, 2, and 4 of MRJP1 to contribute to the antimicrobial properties of honey. The detection of defensin-1 in all honeys indicates that honeybees supply this antimicrobial peptide to all honeys similarly to MRJP1. Detection of hymenoptaecin in only three honeys (H03, H08, and H13) indicates that addition of this antimicrobial polypeptide that suppresses both G+ and G− bacteria in the honeybee⁴² to honey can be context-dependent. The reasons for the variability of hymenoptaecin abundance warrant future investigation; we hypothesize that these reasons relate to the microbial pressures in honeybee colonies and/or entire sites.

The detection of the antimicrobial peptides defensin-1 and hymenoptaecin using a proteomic LFQ nLC-MS/MS technique is important for future investigations. The discovery of hymenoptaecin in honey is one of the key findings that should help increase our understanding of its unique antimicrobial properties.

Glucose Dehydrogenases as Supplementary to Glucose Oxidase Function? Another important factor that contributes to the antimicrobial activity of honey is the ability to maintain H₂O₂ through oxidation by glucose oxidase (GOx; also known as notatin), which is an important enzyme of the honeybee hypopharyngeal glands.⁴³ A previous study using the immunoblot technique with an anti-honeybee GOx antibody showed high variability among honey samples.⁴³ Our results show a relatively constant proportion of this enzyme in the proteome, but the differences in the total protein content among the honeys (Table 2) necessarily leads to natural quantitative GOx differences in honeys. The principal function of GOx is to mediate the conversion of β-D-glucose to D-glucono-1,5-lactone (GDL), which is further nonenzymatically hydrolyzed to gluconic acid (GA),^{44,45} a major acid in most honeys.^{46,47} Importantly, however, GDL can also be produced by the identified isoforms of Gdh.⁴⁸ Both GOx and Gdhs are utilized in biotechnology for detecting glucose, and the important difference is that GOx can utilize O₂ as an electron acceptor, thereby producing H₂O₂, while Gdhs utilize different electron acceptors.⁴⁸ After an individual inspection of the proteins identified, we report the detection of four Gdhs produced by different genes, and while all four Gdhs have the

GMC_oxred_C superfamily domain, one has an additional PLN02985 superfamily domain for squalene monooxygenase.

One of the enzyme activities found in honey is that of catalase, which is thought to originate from pollen, nectar, and microorganisms.²² We failed to detect catalase or peroxidase that could be assigned to the honeybee taxonomy, which supports the hypothesis of different sources, unlike the honeybee GOx, which produces H₂O₂.

These findings suggest that in addition to the well-known presence of GOx we identified Gdh isoforms as present in the honey. Although Gdhs cannot produce H₂O₂ as GOx does, they can contribute to the formation of GA in the absence of O₂, while GOx requires O₂. The discovery of Gdhs in honey will contribute to clarification of the formation of GA and GDL, which are important for maintaining honey acidity and contribute greatly to taste, stability, flavor, and pharmaceutical value.⁴⁶

Venom-like Proteins and Allergens in Honey. In addition to its negative allergenic effects, honeybee venom has a very important role in medicine because it is used to treat various diseases, e.g., arthritis,⁴⁹ cancer,⁵⁰ and skin diseases.⁵¹ Furthermore, bee venom has been shown to suppress microbes that cause diseases, e.g., bacterial agents such as *Borrelia burgdorferi*⁵² and viruses.⁵³ Thus, the proteins in honey that are normally known to occur in venom are of exceptional interest. First, the current allergen spectrum of honeybee contains 12 confirmed allergens listed on the WHO/IUIS list of allergens,⁵⁴ and all the honeys in our study contained the allergens Api m 10 (icarapin) and Api m 11 (MRJP9 was specifically identified; note that according to the official WHO/IUIS GenBank IDs NP_001011564.1 and AAY21180.1,⁵⁴ Api m 11 is two MRJPs, 8 and 9, respectively). Moreover, in eight honeys, we detected the allergen Api m 2 (hyaluronidase). Furthermore, our analysis revealed the presence of analogues to Api m 7 (CUB serine protease) and Api m 3 (acid phosphatase) across the tested honeys. Notably, the two identified isoforms of venom acid phosphatase Acph-1-like have 83% identity and 100% query coverage to each other and 36% and 37% identity (both 93% identity) to Api m 3 (GenBank: AAYS7281.1), respectively. Finally, Api m 1 (phospholipase A2), melittin (Api m 4), and Api m 12 (vitellogenin) were detected in only one honey/analysis (Table S1, Supporting Information); thus, these trace identifications were eliminated from the final list of 71 protein hits. In addition, proteomic studies have shown that honeybee venom proteins do not include only venom allergens;^{55–57} therefore, we compared our data set with protein identifications from previous proteomic studies (see Table 1). This step revealed that in honey we identified additional proteins previously found in honeybee venom. While CLIP/venom serine protease Bi-VSP (Bi-VSP), CUB/venom serine protease 34, venom acid phosphatase Acph-1-like (GenBank: XP_006569974.1), Chitinase-like protein Idgf4, transferrin 1, and glucosylceramidase-like were detected in all honeys, apolipoproteins, Hsc70-3/Hsp70Ab, histone h4, ubiquitin-60S ribosomal protein L40, elongation factor 1-α, Obp 14, and hexamerin 70a were detected in only some of the honeys. Furthermore, an ovalbumin-related protein X/serine protease inhibitor 88Ea (antithrombin III) was detected in 12 honeys.

These exact matches to honeybee venom proteins, especially those that are confirmed allergens, are important in relation to possible honey allergies as well as honey's potential to remedy these allergies. Further experiments are necessary to show

whether there are allergenic proteins in honey in addition to those that we identified and that are listed on the WHO/IUIS list.⁵⁴

Proteases and Protease Inhibitors in Honey. As noted in the previous section, we identified the CUB serine protease (venom serine protease 34), which is an analogue to the Api m 7 allergen. Furthermore, we identified the Bi-VSP, which is an analogue in *Bombus ignitus* that has been shown to be an important component of venom and to have a dual function as an arthropod phenoloxidase (PO) cascade triggering factor and as a fibrin(ogen)olytic enzyme in mammals.⁵⁸ The analysis of CCDs revealed that the honeybee Bi-VSP has the two domains CLIP and Tryp_SPC, and its domain architecture is therefore similar to, for example, *Drosophila melanogaster* hayn (GenBank: NP_573296) and *Bombyx mori* prophenoloxidase activating factor 3 (GenBank: AAL31707). Of these, hayn has been described as a component of the systemic wound response (SWR) and, after activation by local wounding, converts pro-phenoloxidase (PPO) to PO.⁵⁹ Furthermore, we identified a carboxypeptidase Q, which is structurally completely different (identity and CCDs: M28_Pgcp_like and PA_M28_2) from the allergenic serine carboxypeptidase Api m 9, which has the peptidase S10 domain. Finally, we identified the protease assigned as PO-activating factor 2. The proteases in honey were investigated in a previous study,²⁶ in which the authors identified serine protease activities as predominant, but in that study, the particular proteases were not on the MS identification list of 2DE spots.

Based on our identification, the serine protease activity in honey is likely regulated by a serine protease inhibitor. One inhibitor has the serpin domain and is assigned as serine protease inhibitor 88Ea. Furthermore, we identified three chymotrypsin inhibitors, which are products of different genes and with the same domain TIL. Identification of the protease inhibitors in honey is a very interesting finding and warrants future investigations on their function.

Herein, we identified the proteases and protease inhibitors in honey; however, their exact roles in honey remain unclear. The proteases in honey can degrade honey proteins, especially MRJP1.²⁶ Furthermore, it can be suggested that the proteases that play a role in the PO cascade have potential to trigger the honeybee immune system. Their role in relation to humans as consumers is as possible allergens but also in allergy prevention. The serine protease inhibitors can inhibit the function of honey proteases; however, we suggest their potential involvement in antimicrobial activity.⁶⁰

Royal Jelly Related Proteins in Honey. Compared to the low protein content in honey,¹⁵ the proteins in royal jelly can reach up to approximately 50% dry matter, and MRJPs are of highest abundance.⁶¹ Previous studies have shown that MRJPs are the proteins with the greatest abundance in honey; however, these identifications were restricted to MRJPs1–5.^{25–27} Evaluation of the proteomic results of this study indicated that eight MRJPs were identified, and MRJP8 was the missing MRJP. Notably, according to the GenBank IDs at the WHO/IUIS,⁵⁴ the allergen Api m 11 includes MRJP9 (GenBank: AAY21180.1) and MRJP8 (GenBank: NP_001011564.1), which have 63% identity and 96% query coverage to each other. The overall quantitative data intensity (clearly illustrated in heatmaps in Figure 1) and MS/MS counts (Table 1) show that MRJP1 has the highest abundance, consistent with earlier research.²⁵ The MRJPs can be divided into groups according to their decreasing relative abundance

levels/intensities in honey: (i) MRJP1; (ii) MRJP2, MRJP3, and MRJP5; (iii) MRJP7, MRJP4, and MRJP6; and (iv) MRJP9. The detection of MRJPs in this study is broader and more complete than the detection of these proteins (also considering isoforms related to protein IDs) in royal jelly alone in previous studies (Table 1; see refs 62–70, but the referenced studies confirm that these eight MRJPs are found in royal jelly). Royal jelly also contains a set of other proteins that we identified in honey and that were previously identified in royal jelly (Table 1). For instance, the antimicrobial peptides defensin-1^{69,70} and hymenoptaecin⁶⁹ were detected in royal jelly. Furthermore, we emphasize that various proteases and the proven allergens Api m 10 and 11 are shared across venom, royal jelly, and honey.

Overall, we detected in honeys the major proteins that were identified in royal jelly in previous studies. Notably, however, the variety of proteins identified here matched the results of various proteomic studies of royal jelly showing relatively high variation in identification. Our quantitative comparison indicated that the MRJPs and other royal jelly related proteins were in most cases constantly present in the tested honeys.

Amylase and α -Glucosidase in Honey. The important protein components of honeys are considered to be enzymes, such as amylase and α -glucosidase (invertase). Invertase was the subject of studies approximately a century ago.⁷¹ Three α -glucosidases with differences in substrate specificity have been characterized from honeybees.^{72–74} A localization study showed that α -glucosidase III (HBG3) is secreted from the hypopharyngeal glands and should therefore be the main honey α -glucosidase.²¹ Herein, we report two results corresponding to HBG3 α -glucosidase with 99% identity. The results indicate that HBG3 is a highly abundant protein in honey. Furthermore, we detected α -glucosidase II (HBG2) in one honey/analysis, and the result is on the list of discarded hits (Table S1 Supporting Information). HBG2 has been previously located in the honeybee ventriculus and hemolymph.²¹ Thus, HBG2 can be considered a rare protein in honey that can originate from the ventriculus. Amylase is similar to HBG3 produced from the hypopharyngeal glands.¹⁹ Despite uncertainty, amylase activity in honey has become a parameter used to evaluate the freshness of honey,⁷⁵ and importantly, amylase can potentially be added by producers to mask syrup adulteration of honey.⁷⁶ In this study, we analyzed only authentic honeys and identified honeybee α -amylase as two isoforms with 99% identity; one less abundant isoform was not detected in three honeys. In addition, the protein hit of amylase that was detected in all the samples analyzed showed very low proteomic variation; again, however, the varying total protein contents (Table 2) in honeys should produce different amylase abundance levels or enzyme activities.

These results indicated that honeys contain honeybee α -glucosidase HBG3 and α -amylase at relatively steady and high levels. In previous proteomic studies, Rossano et al.²⁶ and Di Girolamo et al.²⁷ identified α -glucosidase but not α -amylase; however, α -amylase was detected by LC-MS/MS in Borutinskaite et al.³⁰ and Zhang et al.²⁹ This difference can be explained by the lower abundance of α -amylase as indicated by the intensity and spectral counts reported in this study.

Other Proteins. Furthermore, we identified other proteins that were in all honeys and additional proteins that were lacking in some honeys. For these proteins as well as the other proteins identified, it is appropriate to consider quantitative values, such as intensity and/or MS/MS counts. Considering

the possibility that honey can be contaminated to different extents by honeybee proteins that are not supplied to the honey by workers is also important; thus, the proteins might belong to, for example, larvae remaining in the combs at the time of honey extraction. Furthermore, the predicted presence of signal protein in the protein, as indicated in Table 1, may suggest excretion.

As noted in the introduction, compared to this study, previous MS-based proteomic studies have reported only a few honeybee-related proteins in honey. In addition to the proteins described above, we highlight that all the honeys contained laccase-5; this protein is interesting because laccases can act on various substrates, including phenols,⁷⁷ which are abundant in honey.⁷⁸ Next, a group of esterases/lipases, i.e., esterase B1, lipase member H-A, and pancreatic triacylglycerol lipases, were also abundant in honeys. These proteins may participate in modification of the lipid components in honey.⁷⁹ Other proteins listed in Table 1 and Table S1, Supporting Information, are also worthy of our attention (e.g., the four hexamerins; uncharacterized protein LOC413627, which has two SGL domains and is similar to regucalcin according to Blastp; two glucosylceramidases that should participate in sphingolipid metabolism; xanthine dehydrogenase, which likely participates in purine metabolism; and lysosomal alpha-mannosidase).

STRING Analysis of Honey Proteins. To reveal the possible connections among the variety of proteins identified, we performed a STRING^{80,81} analysis. Gene names were determined (Table 1 and Table S1, Supporting Information) for the protein results, and this list was input into STRING^{80,81} to produce the network, which is shown in Figure S2; Table S2, Supporting Information, provides the results of the analyses. The functional enrichments in the network confirm that one of the key biological processes is the defense response, including defense response to bacteria, carbohydrate metabolic process, organic substrate metabolic process, and innate immune response.

Complementary Sugar Analysis. Sugars—mainly glucose and fructose—rather than proteins are the main components of honey.¹⁴ A complementary sugar analysis of the tested honeys using the HPLC-ELSD technique showed that these honeys contained glucose and fructose; other sugars were not detected. Table 2 shows the concentration of sugars and the fructose/glucose (fru/glu) ratio in each honey. The proteomic results indicated a difference in the eucalyptus honey (Figure 1, Table 1) from the remaining 12 honeys, and an analogous situation was found in the fru/glu ratio, which was highest in the eucalyptus honey. Finally, as noted above, the eucalyptus honey had a relatively high total protein content in our study and in a previous report.⁸⁶ Thus, the proteomic analysis, total protein content, and high fru/glu ratio all indicate the complex dissimilarity of eucalyptus honey.

■ EXPERIMENTAL SECTION

Experimental Design. Thirteen authentic honeys from verified sources were used in this study. Twelve honeys originated from Czechia, and the eucalyptus honey was from Galicia, Spain. With the exception of two honeys (H08 and H10), which were from our own production (Crop Research Institute, Prague-Ruzyne), all the remaining honeys were from different sites. The honey denoted as the honey of the year in Czechia was a honey from Kosire in inner Prague. The honey sample denoted as wild bees was produced by honeybees found below the roof of a house in Prague-Ruzyne. Each sample cleaned using the gel filtration technique was analyzed in three

replicates using the label-free proteomic approach. All of the samples digested with trypsin were nLC-MS/MS (Thermo) analyzed in three biological and three analytical replicates in one analytical series; thus, 39 nLC-MS runs were performed consecutively.

Sample Preparation. The honey samples were dissolved in 0.2- μ m-filtered Nanopure water (Barnstead, Thermo) at a ratio of 1 g of honey to 2 mL of H₂O. Then, the samples were cleaned using PD MidiTrap G-25 columns (Cat. No. 28-9180-08, GE Healthcare Life Sciences). These procedures were performed on ice to prevent degradation. Sample aliquots were concentrated by lyophilization of the frozen contents in 15 mL centrifuge tubes covered with filters in a PowerDry LL3000 (Thermo). Then, the protein content was determined using Bradford reagent (Sigma-Aldrich) in a microplate, and bovine serum albumin was used as a calibration standard. The protein determination was performed in three lyophilized protein aliquots, and two analytical replicates were performed. The proteins determined were also used for the calculation of the total proteins in the honey. Furthermore, the protein profile was tested using 1D SDS-PAGE, and samples of three separate aliquots for each honey were used for nLC-MS/MS analysis.

Proteomic Analysis of Honey. The lyophilizate was reconstituted in 100 mM triethylammonium bicarbonate buffer (1.0 M, pH = 8.5 $\pm$ 0.1) containing 2% sodium deoxycholate (SDC) and boiled at 95 °C for 5 min. Protein concentration was determined using a BCA protein assay kit (Thermo), and 25 μ g of protein per sample was used for MS sample preparation. Cysteines were reduced with a 5 mM final concentration of tris(2-carboxyethyl)phosphine hydrochloride (60 °C for 60 min) and blocked with a 10 mM final concentration of methylmethanethiosulfonate (10 min at room temperature). Samples were digested with trypsin (trypsin/protein ratio 1/20) at 37 °C overnight. Then, the samples were acidified with trifluoroacetic acid to a final concentration of 1%. SDC was removed by extraction to EtOAc,⁸² and peptides were desalted on a Michrom C18 column. Dried peptides were resuspended in 25 μ L of H₂O containing 2% acetonitrile and 0.1% trifluoroacetic acid. For analysis, 12 μ L of sample was injected.

A reversed-phase nanocolumn (EASY-Spray column, 50 cm $\times$ 75 μ m ID, PepMap C18, 2 μ m particles, 100 Å pore size) was used for LC-MS/MS analysis. Mobile phase buffer A was composed of water and 0.1% formic acid (FA). Mobile phase B was composed of CH₃CN and 0.1% FA. Samples were loaded onto the trap column (Acclaim PepMap300, C18, 5 μ m, 300 Å wide pore, 300 μ m $\times$ 5 mm) at a flow rate of 15 μ L/min. Loading buffer was composed of H₂O, 2% CH₃CN, and 0.1% trifluoroacetic acid. Peptides were eluted with gradient of B from 4% to 35% over 60 min at a flow rate of 300 nL/min. Eluting peptide cations were converted to gas-phase ions by electrospray ionization and analyzed using nLC-MS/MS. Survey scans of peptide precursors from 350 to 1400 m/z were performed at 120 K resolution (at 200 m/z) with a 5×10^5 ion count target. Tandem MS was performed by isolation at 1.5 Th with the quadrupole, HCD fragmentation with normalized collision energy of 30, and rapid scan MS analysis in the ion trap. The MS 2 ion count target was set to 10^4 , and the max injection time was 35 ms. Only those precursors with charge state 2–6 were sampled for MS². The dynamic exclusion duration was set to 45 s with a 10 ppm tolerance around the selected precursor and its isotopes. Monoisotopic precursor selection was turned on. The instrument was run in top speed mode with 2 s cycles.^{83–85}

Data Evaluation. The data were evaluated using the v1.6.3.3 MaxQuant LFQ algorithm^{86,87} and the *A. mellifera* taxonomy (txid7460) database downloaded from NCBI on October 29, 2018, which consisted of 31 507 sequences. The false discovery rate was set to 1% for both proteins and peptides, and we specified a minimum length of seven amino acids. Methyl/Thio was selected as a fixed modification, and N-terminal protein acetylation and methionine oxidation were selected as variable modifications. Our created database contained the gi numbers for easier comparison with previous studies. For the details on MaxQuant parameters and database used, in the data availability section, we provide the entire

combined folder and database for download together with the raw data.

The data were further evaluated in v1.6.2.2 Perseus.³⁵ Some of the evaluation details are also explained in the [Results and Discussion section](#). The LFQ intensity was log 2-transformed, and the data set was filtered of the contaminants and reverse hits. Further, the matrix was filtered to proteins that contained at least three valid values in total. Inspection of the histograms (Figure S3, [Supporting Information](#)) indicated a normal distribution of the data. The data were averaged by the median. We found that heatmap presentation is appropriate for the data evaluation. Thus, we generated two heatmaps that illustrate the proteome differences. One heatmap resulted after the missing values were replaced by the constant “0”, and the second heatmap was generated after the missing values were replaced based on a normal distribution (width: 0.3; downshift: 1.5). Euclidian clustering was performed (number of clusters: 300; max. interactions: 10; restarts: 1).

In the data evaluation, we verified protein isoforms manually using Blastp.⁸⁸ The protein names were curated individually. We also verified the obsolete protein identification numbers provided in the studies related to honey, venom, and royal jelly proteomics through NCBI-Protein and Blastp.⁸⁸ Furthermore, we additionally verified the protein names that may change in time, moreover if the headers “uncharacterized” were provided (for example, in one study,³⁰ we considered a previously unreported but identified protein in honey). However, such consideration complicates the fact that some studies did not provide sufficient identification details, including peptides or raw data.

Furthermore, we searched the proteins using SignalP 4.1,³² Signal-CP,³³ and Signal-3L 2.0³⁴ to predict secretion through the presence of signal peptides. The first search for the prediction of signal peptides was performed via SignalP 4.1,³² and we also verified the presence of signal peptides in the UniProt. We ensured that the sequence analyzed was complete rather than partial. Note that in some proteins Signal-CP³³ or Signal-3L 2.0³⁴ was necessary to complement identification of the signal peptide. Finally, we performed the functional protein association network analysis of the 71 protein hits detected in honey using STRING v 10.5.^{80,81}

Sugar Analysis. For sugar analysis, samples of honey were prepared in 1:4 CH₃CN/H₂O at a concentration of 40 mg/mL. A 1 mL amount of the sample was filtered through regenerated cellulose syringe filters of 13 mm diameter and 0.45 µm pore size (TR-200435; Teknokroma) into a glass vial. The analysis was conducted using an Agilent 1200 series HPLC (Agilent) equipped with a degasser, quaternary pump, autosampler, thermostat, and 1260 Infinity ELSD detector. The chromatographic separation was performed on a Zorbax NH₂ instrument (150 mm × 4.6 mm, 5 µm; Agilent) protected by a high-pressure guard column with an NH₂ cartridge (4.6 mm × 12.5 mm, 5 µm; Agilent). The column temperature was maintained at 23 °C throughout the analysis. The ELSD parameters were a gas flow rate of 1.2 SLM, evaporator 30 °C, and nebulizer 30 °C. Separation was performed with a flow rate of 1 mL/min using the isocratic elution with 4:1 CH₃CN/H₂O and an injection volume of 10 µL. The instrument control and data evaluation were performed using OpenLab software (Agilent). The calibration and extraction efficiency were determined with glucose (Cat. No. G8270, Sigma-Aldrich) and fructose (Cat. No. F0127, Sigma-Aldrich).

Data Availability. The accession numbers for the raw nLC-MS/MS runs reported in this paper are MassIVE MSV000083096 (<https://doi.org/10.25345/c5xk5n>) and PXD011569. Furthermore, we provide for download the entire “combined” folder of MaxQuant (v1.6.3.3) data processing, the database used for the search, and the MaxQuant parameters.

■ ASSOCIATED CONTENT

● Supporting Information

The Supporting Information is available free of charge on the [ACS Publications website](#) at DOI: [10.1021/acs.jnatprod.8b00968](https://doi.org/10.1021/acs.jnatprod.8b00968).

Proteomic data evaluation details; the STRING analysis; sample chromatograms of the sugar analysis; the STRING network; histograms of the 39 nLC-MS/MS runs ([PDF](#))

■ AUTHOR INFORMATION

Corresponding Author

*E-mail (T. Erban): arachnid@centrum.cz.

ORCID

Tomas Erban: 0000-0003-1730-779X

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study was supported by Grant No. QK1820088 of the Ministry of Agriculture of the Czech Republic. We thank M. Markovic and M. Simonovsky for their valuable help.

■ REFERENCES

- (1) Schmidt, J. O. In *Bee Products*; Mizrahi, A.; Lensky, Y., Eds.; Springer: Boston, MA, 1997; pp 15–26.
- (2) Nayik, G. A.; Shah, T. R.; Muzaffar, K.; Wani, S. A.; Gull, A.; Majid, I.; Bhat, F. M. *Universal J. Pharm.* **2014**, *3*, 5–8.
- (3) Molan, P. C. *Bee World* **1992**, *73*, 5–28.
- (4) Alvarez-Suarez, J. M.; Tulipani, S.; Romandini, S.; Bertoli, E.; Battino, M. *Mediterr. J. Nutr. Metab.* **2010**, *3*, 15–23.
- (5) Saranraj, P.; Sivasakthi, S.; Feliciano, G. D. *Adv. Biol. Res.* **2016**, *10*, 271–289.
- (6) The Lancet. *Lancet* **2018**, *392*, 712–712.
- (7) Martinotti, S.; Calabrese, G.; Ranzato, E. *Mol. Cell. Biochem.* **2017**, *435*, 185–192.
- (8) Ranzato, E.; Martinotti, S. *Cellular and Molecular Mechanisms of Honey Wound Healing*; Nova Science Publishers: Hauppauge, NY, 2014.
- (9) Molan, P. C. *Am. J. Clin. Dermatol.* **2001**, *2*, 13–19.
- (10) Hidaka, S.; Okamoto, Y.; Ishiyama, K.; Hashimoto, K. *J. Periodontal Res.* **2008**, *43*, 450–458.
- (11) Ahmed, S.; Othman, N. H. *Evid. Based Complement. Alternat. Med.* **2013**, *2013*, 829070.
- (12) Othman, N. H. *J. Tradit. Complement. Med.* **2012**, *2*, 276–283.
- (13) EC. Report from the Commission to the European Parliament and the Council on the implementation of the measures concerning the apiculture sector of Regulation (EU) No 1308/2013 of the European Parliament and of the Council establishing a common organisation of the markets in agricultural products, COM/2016/0776 final, Document 52016DC0776, December 7, 2016. European Commission (EC): Brussels, 2016. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016DC0776> (accessed October 22, 2018).
- (14) Soares, S.; Amaral, J. S.; Oliveira, M. B. P. P.; Mafra, I. *Compr. Rev. Food Sci. Food Saf.* **2017**, *16*, 1072–1100.
- (15) Chua, L. S.; Lee, J. Y.; Chan, G. F. *Anal. Bioanal. Chem.* **2013**, *405*, 3063–3074.
- (16) Won, S.-R.; Li, C.-Y.; Kim, J.-W.; Rhee, H.-I. *Food Chem.* **2009**, *113*, 1334–1338.
- (17) Afroz, R.; Tanvir, E. M.; Zheng, W.; Little, P. J. *J. Clin. Exp. Pharmacol.* **2016**, *6*, 3.
- (18) Babacan, S.; Rand, A. G. *J. Food Sci.* **2005**, *70*, C413–C418.
- (19) Ohashi, K.; Natori, S.; Kubo, T. *Eur. J. Biochem.* **1999**, *265*, 127–133.
- (20) Costa, R. A. C.; da Cruz-Landim, C. *Genet. Mol. Res.* **2005**, *4*, 616–623.
- (21) Kubota, M.; Tsuji, M.; Nishimoto, M.; Wongchawalit, J.; Okuyama, M.; Mori, H.; Matsui, H.; Surarit, R.; Svasti, J.; Kimura, A.; Chiba, S. *Biosci., Biotechnol., Biochem.* **2004**, *68*, 2346–2352.
- (22) Kwakman, P. H. S.; Zaat, S. A. J. *IUBMB Life* **2012**, *64*, 48–55.

- (23) Lee, D.-C.; Lee, S.-Y.; Cha, S.-H.; Choi, Y.-S.; Rhee, H.-I. *Korean J. Food Sci. Technol.* **1998**, *30*, 1–5 (in Korean with English abstract) .
- (24) Marshall, T.; Williams, K. M. *Anal. Biochem.* **1987**, *167*, 301–303.
- (25) Won, S.-R.; Lee, D.-C.; Ko, S. H.; Kim, J.-W.; Rhee, H.-I. *Food Res. Int.* **2008**, *41*, 952–956.
- (26) Rossano, R.; Larocca, M.; Polito, T.; Perna, A. M.; Padula, M. C.; Martelli, G.; Riccio, P. *PLoS One* **2012**, *7*, e49164.
- (27) Di Girolamo, F.; D'Amato, A.; Righetti, P. G. *J. Proteomics* **2012**, *75*, 3688–3693.
- (28) Chua, L. S.; Lee, J. Y.; Chan, G. F. *Anal. Lett.* **2015**, *48*, 697–709.
- (29) Zhang, Y.-Z.; Chen, Y.-F.; Wu, Y.-Q.; Si, J.-J.; Zhang, C.-P.; Zheng, H.-Q.; Hu, F.-L. *Food Res. Int.* **2019**, *116*, 362–369.
- (30) Borutinskaite, V.; Treigyte, G.; Ceksteryte, V.; Kurtinaitiene, B.; Navakauskiene, R. *J. Food Nutr. Res.* **2018**, *57*, 57–69.
- (31) Azevedo, M. S.; Valentim-Neto, P. A.; Seraglio, S. K. T.; da Luz, C. F. P.; Arisi, A. C. M.; Costa, A. C. O. *J. Sci. Food Agric.* **2017**, *97*, 4515–4519.
- (32) Petersen, T. N.; Brunak, S.; von Heijne, G.; Nielsen, H. *Nat. Methods* **2011**, *8*, 785–786.
- (33) Chou, K.-C.; Shen, H.-B. *Biochem. Biophys. Res. Commun.* **2007**, *357*, 633–640.
- (34) Zhang, Y.-Z.; Shen, H.-B. *J. Chem. Inf. Model.* **2017**, *57*, 988–999.
- (35) Tyanova, S.; Temu, T.; Sinitsyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J. *Nat. Methods* **2016**, *13*, 731–740.
- (36) Bauer, L.; Kohlich, A.; Hirschwehr, R.; Siemann, U.; Ebner, H.; Scheiner, O.; Kraft, D.; Ebner, C. *J. Allergy Clin. Immunol.* **1996**, *97*, 65–73.
- (37) Helbling, A.; Peter, C.; Berchtold, E.; Bogdanov, S.; Muller, U. *Allergy* **1992**, *47*, 41–49.
- (38) Marchler-Bauer, A.; Bo, Y.; Han, L.; He, J.; Lanczycki, C. J.; Lu, S.; Chitsaz, F.; Derbyshire, M. K.; Geer, R. C.; Gonzales, N. R.; Gwadz, M.; Hurwitz, D. I.; Lu, F.; Marchler, G. H.; Song, J. S.; Thanki, N.; Wang, Z.; Yamashita, R. A.; Zhang, D.; Zheng, C.; Geer, L. Y.; Bryant, S. H. *Nucleic Acids Res.* **2017**, *45*, D200–D203.
- (39) Geer, L. Y.; Domrachev, M.; Lipman, D. J.; Bryant, S. H. *Genome Res.* **2002**, *12*, 1619–1623.
- (40) Kwakman, P. H. S.; te Velde, A. A.; de Boer, L.; Speijer, D.; Vandenbroucke-Grauls, C. M. J. E.; Zaai, S. A. J. *FASEB J.* **2010**, *24*, 2576–2582.
- (41) Brudzynski, K.; Sjaarda, C. *PLoS One* **2015**, *10*, e0120238.
- (42) Casteels, P.; Ampe, C.; Jacobs, F.; Tempst, P. *J. Biol. Chem.* **1993**, *268*, 7044–7054.
- (43) Bucekova, M.; Valachova, I.; Kohutova, L.; Prochazka, E.; Klaudiny, J.; Majtan, J. *Naturwissenschaften* **2014**, *101*, 661–670.
- (44) White, J. W., Jr.; Subers, M. H.; Schepartz, A. I. *Biochim. Biophys. Acta, Spec. Sect. Enzymol. Subj.* **1963**, *73*, 57–70.
- (45) Bankar, S. B.; Bule, M. V.; Singhal, R. S.; Ananthanarayan, L. *Biotechnol. Adv.* **2009**, *27*, 489–501.
- (46) Ramachandran, S.; Fontanille, P.; Pandey, A.; Larroche, C. *Food Technol. Biotechnol.* **2006**, *44*, 185–195.
- (47) Mato, L.; Huidobro, J. F.; Simal-Lozano, J.; Sancho, M. T. *J. Food Prot.* **2003**, *66*, 2371–2376.
- (48) Ferri, S.; Kojima, K.; Sode, K. *J. Diabetes Sci. Technol.* **2011**, *5*, 1068–1076.
- (49) Kang, S. S.; Pak, S. C.; Choi, S. H. *Am. J. Chin. Med.* **2002**, *30*, 73–80.
- (50) Orsolic, N. *Cancer Metastasis Rev.* **2012**, *31*, 173–194.
- (51) Senel, E.; Kuyucu, M.; Suslu, I. *Anc. Sci. Life* **2014**, *33*, 192–193.
- (52) Socarras, K. M.; Theophilus, P. A. S.; Torres, J. P.; Gupta, K.; Sapi, E. *Antibiotics (Basel)* **2017**, *6*, E31.
- (53) Uddin, M. B.; Lee, B.-H.; Nikapitiya, C.; Kim, J.-H.; Kim, T.-H.; Lee, H.-C.; Kim, C. G.; Lee, J.-S.; Kim, C.-J. *J. Microbiol.* **2016**, *54*, 853–866.
- (54) WHO/IUIS. *Allergen Nomenclature*. WHO/IUIS Allergen Nomenclature Sub-Committee; 2019. <http://www.allergen.org/> (accessed March 27, 2019).
- (55) Van Vaerenbergh, M.; Debyser, G.; Devreese, B.; de Graaf, D. C. *J. Proteomics* **2014**, *99*, 169–178.
- (56) Danneels, E. L.; Van Vaerenbergh, M.; Debyser, G.; Devreese, B.; de Graaf, D. C. *Toxins* **2015**, *7*, 4468–4483.
- (57) Li, R.; Zhang, L.; Fang, Y.; Han, B.; Lu, X.; Zhou, T.; Feng, M.; Li, J. *BMC Genomics* **2013**, *14*, 766.
- (58) Choo, Y. M.; Lee, K. S.; Yoon, H. J.; Kim, B. Y.; Sohn, M. R.; Roh, J. Y.; Je, Y. H.; Kim, N. J.; Kim, I.; Woo, S. D.; Sohn, H. D.; Jin, B. R. *PLoS One* **2010**, *5*, e10393.
- (59) Nam, H.-J.; Jang, I.-H.; You, H.; Lee, K.-A.; Lee, W.-J. *EMBO J.* **2012**, *31*, 1253–1265.
- (60) Mandal, S. M.; Porto, W. F.; De, D.; Phule, A.; Korpole, S.; Ghosh, A. K.; Roy, S. K.; Franco, O. L. *Analyst* **2014**, *139*, 464–472.
- (61) Fratini, F.; Cilia, G.; Mancini, S.; Felicioli, A. *Microbiol. Res.* **2016**, *192*, 130–141.
- (62) Yu, F.; Mao, F.; Jianke, L. *J. Proteome Res.* **2010**, *9*, 2207–2215.
- (63) Fujita, T.; Kozuka-Hata, H.; Ao-Kondo, H.; Kunieda, T.; Oyama, M.; Kubo, T. *J. Proteome Res.* **2013**, *12*, 404–411.
- (64) Furusawa, T.; Rakwal, R.; Nam, H. W.; Shibato, J.; Agrawal, G. K.; Kim, Y. S.; Ogawa, Y.; Yoshida, Y.; Kouzuma, Y.; Masuo, Y.; Yonekura, M. *J. Proteome Res.* **2008**, *7*, 3194–3229.
- (65) Han, B.; Li, C.; Zhang, L.; Fang, Y.; Feng, M.; Li, J. *J. Agric. Food Chem.* **2011**, *59*, 10346–10355.
- (66) Schonleben, S.; Sickmann, A.; Mueller, M. J.; Reinders, J. *Anal. Bioanal. Chem.* **2007**, *389*, 1087–1093.
- (67) Han, B.; Fang, Y.; Feng, M.; Lu, X.; Huo, X.; Meng, L.; Wu, B.; Li, J. *J. Proteome Res.* **2014**, *13*, S928–S943.
- (68) Li, J.; Wang, T.; Zhang, Z.; Pan, Y. *J. Agric. Food Chem.* **2007**, *55*, 8411–8422.
- (69) Zhang, L.; Han, B.; Li, R.; Lu, X.; Nie, A.; Guo, L.; Fang, Y.; Feng, M.; Li, J. *BMC Genomics* **2014**, *15*, 135.
- (70) Feng, M.; Fang, Y.; Han, B.; Xu, X.; Fan, P.; Hao, Y.; Qi, Y.; Hu, H.; Huo, X.; Meng, L.; Wu, B.; Li, J. *J. Proteome Res.* **2015**, *14*, 5327–5340.
- (71) Nelson, J. M.; Cohn, D. J. *J. Biol. Chem.* **1924**, *61*, 193–224.
- (72) Nishimoto, M.; Kubota, M.; Tsuji, M.; Mori, H.; Kimura, A.; Matsui, H.; Chiba, S. *Biosci., Biotechnol., Biochem.* **2001**, *65*, 1610–1616.
- (73) Takewaki, S.-i.; Kimura, A.; Kubota, M.; Chiba, S. *Biosci., Biotechnol., Biochem.* **1993**, *57*, 1508–1513.
- (74) Kimura, A.; Takewaki, S.-i.; Matsui, H.; Kubota, M.; Chiba, S. *J. Biochem.* **1990**, *107*, 762–768.
- (75) Tosi, E.; Martinet, R.; Ortega, M.; Lucero, H.; Re, E. *Food Chem.* **2008**, *106*, 883–887.
- (76) Voldrich, M.; Rajchl, A.; Cizkova, H.; Cuhra, P. *Czech J. Food Sci.* **2009**, *27*, S280–S282.
- (77) Jeon, J.-R.; Baldrian, P.; Murugesan, K.; Chang, Y.-S. *Microb. Biotechnol.* **2012**, *5*, 318–332.
- (78) Pyrzynska, K.; Biesaga, M. *TrAC, Trends Anal. Chem.* **2009**, *28*, 893–902.
- (79) Kapoulas, V. M.; Mastronicolis, S. K.; Galanos, D. S. *Z. Lebensm.-Unters. Forsch.* **1977**, *163*, 96–99.
- (80) Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K. P.; Kuhn, M.; Bork, P.; Jensen, L. J.; von Mering, C. *Nucleic Acids Res.* **2015**, *43*, D447–D452.
- (81) Szklarczyk, D.; Morris, J. H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N. T.; Roth, A.; Bork, P.; Jensen, L. J.; von Mering, C. *Nucleic Acids Res.* **2017**, *45*, D362–D368.
- (82) Masuda, T.; Tomita, M.; Ishihama, Y. *J. Proteome Res.* **2008**, *7*, 731–740.
- (83) Hebert, A. S.; Richards, A. L.; Bailey, D. J.; Ulbrich, A.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J. *Mol. Cell. Proteomics* **2014**, *13*, 339–347.
- (84) Erban, T.; Harant, K.; Chalupnikova, J.; Kocourek, F.; Stara, J. *J. Proteomics* **2017**, *150*, 281–289.

(85) Erban, T.; Sopko, B.; Talacko, P.; Harant, K.; Kadlikova, K.; Halesova, T.; Riddellova, K.; Pekas, A. *J. Proteomics* **2019**, *196*, 69–80.

(86) Cox, J.; Hein, M. Y.; Lubner, C. A.; Paron, I.; Nagaraj, N.; Mann, M. *Mol. Cell. Proteomics* **2014**, *13*, 2513–2526.

(87) Cox, J.; Mann, M. *Nat. Biotechnol.* **2008**, *26*, 1367–1372.

(88) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. *J. Mol. Biol.* **1990**, *215*, 403–410.

Supporting Information

for

The Unique Protein Composition of Honey Revealed by Comprehensive Proteomic Analysis: Allergens, Venom-like Proteins, Antibacterial Properties, Royal Jelly Proteins, Serine Proteases and Their Inhibitors

Tomas Erban,^{†,*} Elena Shcherbachenko,[†] Pavel Talacko,[‡] and Karel Harant[‡]

[†] Proteomics and Metabolomics Laboratory, Crop Research Institute, Drnovska 507/73, Prague 6-Ruzyne, CZ-16106, Czechia

[‡] Proteomics Core Facility, Faculty of Science, Charles University, BIOCEV, Prumyslova 595, Vestec, CZ-25242, Czechia

*arachnid@centrum.cz

This file includes:

- **Table S1.** Proteomic data evaluation details. (next page)
- **Table S2.** STRING analysis.
- **Figure S1.** Sample chromatograms of the sugar analysis.
- **Figure S2.** STRING network.
- **Figure S3.** Histograms of the 39 nLC-MS/MS runs

[illegible]

Table S2. Results of the STRING analysis. (Access date: 3rd November 2018)

Network Stats

number of nodes: 56
number of edges: 14
average node degree: 0.5
avg. local clustering coefficient: 0.229
expected number of edges: 7
PPI enrichment p-value: 0.0136

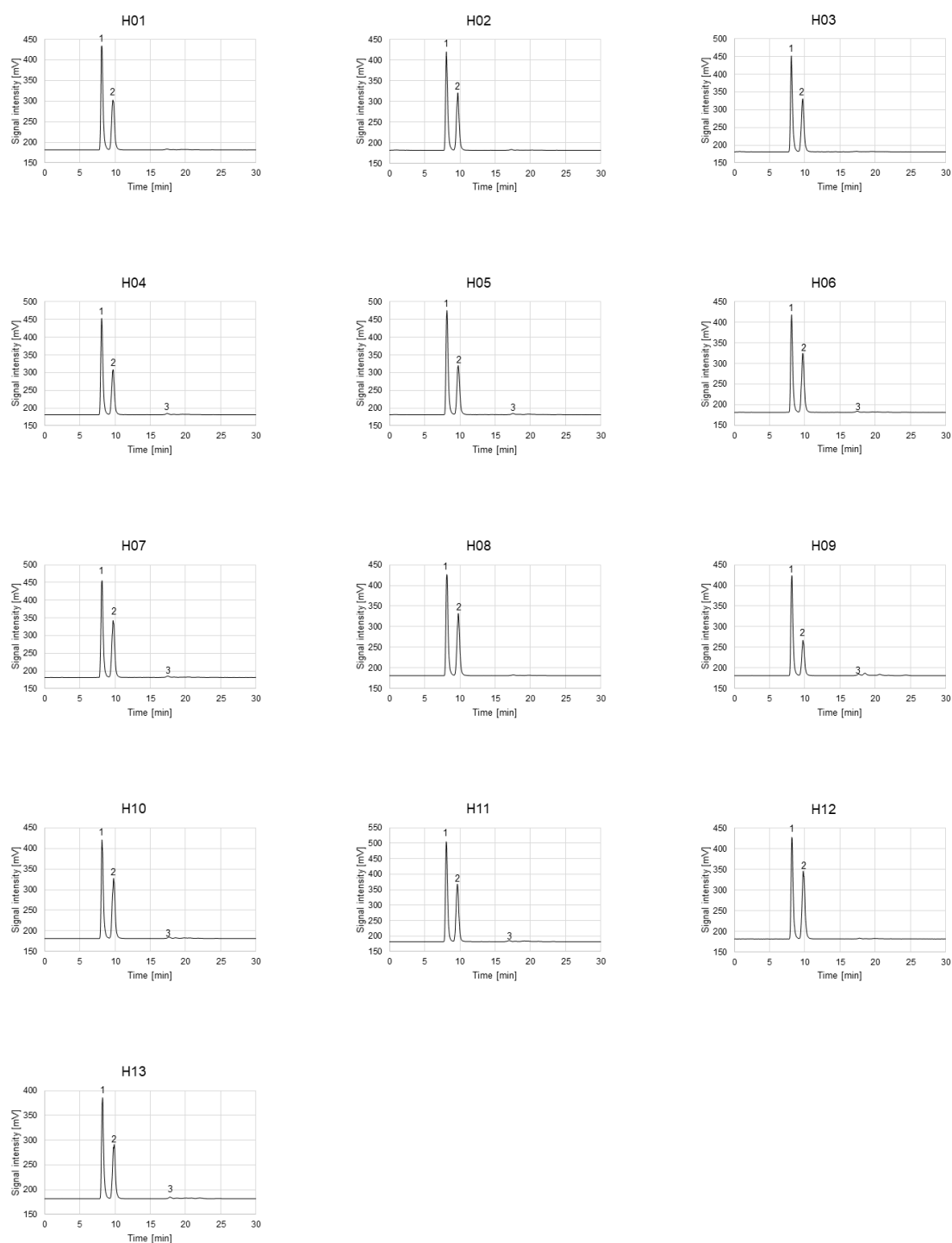
Biological Process (GO)

<i>pathway ID</i>	<i>pathway description</i>	<i>count in gene set</i>	<i>false discovery rate</i>
GO:0008150	biological_process	7	5.98e-06
GO:0006952	defense response	3	0.00247
GO:0042742	defense response to bacterium	3	0.000422
GO:0005975	carbohydrate metabolic process	2	0.00234
GO:0044238	primary metabolic process	3	0.00313
GO:0071704	organic substance metabolic process	3	0.00313
GO:0045087	innate immune response	2	0.0124

Molecular Function (GO)

<i>pathway ID</i>	<i>pathway description</i>	<i>count in gene set</i>	<i>false discovery rate</i>
GO:0003674	molecular_function	4	0.00251
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	2	0.00251
GO:0016787	hydrolase activity	3	0.00251
GO:0043167	ion binding	2	0.0447

Figure S1. Sample chromatograms of the sugar analysis using high-performance liquid chromatography coupled with evaporative light scattering detector (HPLC-ELSD).



Legend: H01–H13 – honey sample; peak No. 1 – fructose; peak No. 2 – glucose; peak No. 3 – indeterminate

Figure S2. The STRING network. (Access date: 3rd November 2018)

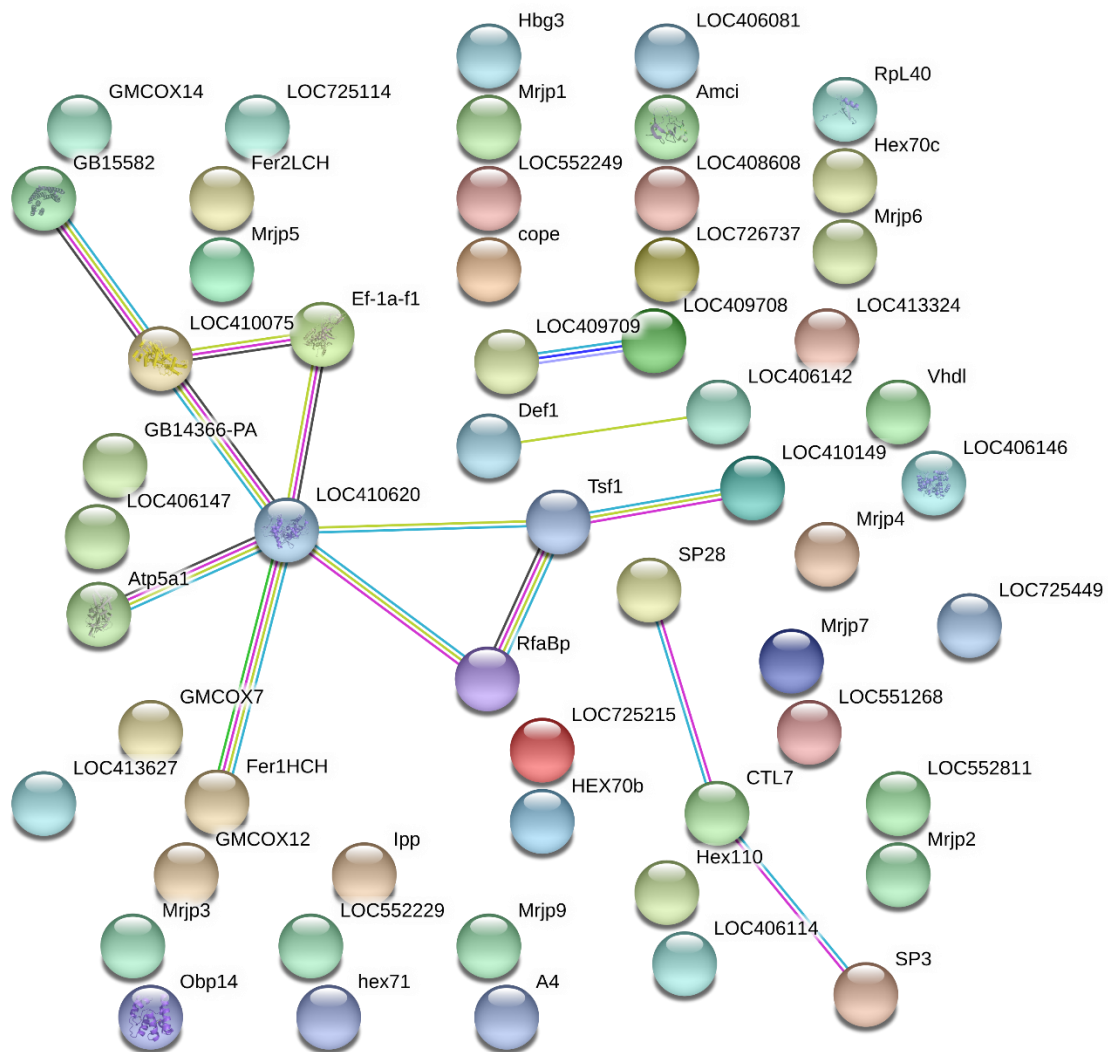


Figure S3. Histograms after of the 39 nLC-MS/MS runs. A) before and B) after the matrix was filtered to proteins that contained at least three valid values in total.

