Received: 7 February 2020 | Revised: 11 March 2020 | Accepted: 22 March 2020

DOI: 10.1111/cea.13604

WILEY

ORIGINAL ARTICLE

Allergens

Measurement of German cockroach allergens and their isoforms in allergen extracts with mass spectrometry

Samuel T. Mindaye[1] | Natalie A. David[1] | Sayyed Amin Zarkesh Esfahani[1] | Coby Schal[2] | Elizabeth C. Matsui[3] | Ronald L. Rabin[1] | Jay E. Slater[1]

[1]Laboratory of Immunobiochemistry, Division of Bacterial, Parasitic, and Allergenic Products, Office of Vaccines Research and Review, CBER/FDA, Silver Spring, MD, USA

[2]Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC, USA

[3]Department of Population Health and Pediatrics, Dell Medical School, The University of Texas at Austin, Austin, TX, USA

Correspondence
Jay E. Slater, Laboratory of Immunobiochemistry, Division of Bacterial, Parasitic, and Allergenic Products, Office of Vaccines Research and Review, CBER/FDA, Silver Spring, MD, USA
Email: Jay.Slater@fda.hhs.gov

Funding information
U.S. Food and Drug Administration

Abstract

Background: Allergen extracts are the primary tool for diagnosis and treatment of allergic diseases. In the United States, most allergen extracts are non-standardized. More sophisticated analytical approaches are needed to characterize these products and enable manufacturers and regulators to better determine potency.

Objective: To expand the multiple reaction monitoring (MRM) assay for an in-depth characterization of German cockroach (GCr; *Blattella germanica*) allergen extracts.

Methods: We applied advanced liquid chromatography (LC) and mass spectrometry (MS) techniques including MRM. The expanded LC/MRM-MS method was optimized to measure known GCr allergens and their isoforms/variants in commercial extracts and environmental samples. We performed isoform-specific allergen measurements in multiple extracts from four commercial sources and extracts prepared using environmental samples from urban homes. To investigate causes of heterogeneity, we examined over 30 extraction process variables.

Results: Evaluation of the commercial extracts confirmed the variability of production lots and commercial sources. Commonly used defatting and extraction protocols yielded extracts with comparable allergen profiles and content. However, the identity and quality of source materials was a major contributor to variability. In comparing commercial GCr extracts to environmental samples, relative quantities of Bla g 1, Bla g 2, Bla g 3, Bla g 4 and Bla g 11 were similar, while Bla g 5, Bla g 6, Bla g 7 and Bla g 8 were present in the environmental samples but largely absent for the commercial extracts.

Conclusions and Clinical Relevance: LC/MRM-MS can be used to measure all known GCr allergens in commercial allergen extracts and environmental samples. Significant differences exist between allergen profiles of commercial extracts and the profiles of environmental samples from dwellings. This analytical platform can serve as a template to achieve better product characterization of similarly complex products.

KEYWORDS

allergens and epitopes, regulatory aspects, venom and insect allergy

1 | INTRODUCTION

Management and therapy of allergic diseases include accurate diagnosis, allergen avoidance, pharmacotherapy and, when appropriate, allergen-specific immunotherapy (AIT)[1] Both diagnosis and AIT require the availability of high-quality species-specific allergen extracts. In the United States, the quality and consistency of commercial allergen extracts are regulated in two ways. For most allergen extracts (non-standardized), manufacturing consistency is the only assurance of a consistent product, as the units used to label these products are not directly associated with extract potency.[2] For the remaining allergen extracts (standardized products), biological potency is confirmed on a lot-by-lot basis and reported in assay-specific units. Potency may be based on total protein, specific allergen content or overall allergenicity. Notably, none of these measures provide a complete content profile of these complex natural products, even though several non-allergen components (beta-glucans, proteases) are known to affect allergic and T-cell responses in vivo.[3,4]

German cockroach (GCr; *Blattella germanica*) is an indoor pest and important source of allergens associated with asthma and allergic rhinoconjunctivitis.[5] Commercial allergen extracts for GCr are non-standardized, and several studies have confirmed low biological potency and a lack of consistency among products manufactured by different companies.[6,7] Further, although 10 GCr allergens have been identified (www.allergen.org), none of them is immunodominant.[8,9]

The FDA identified GCr allergen extracts as an appropriate standardization target in 2001.[10] To achieve this goal, its efforts to measure the allergen content of these products have included ELISA,[6] ID_{50}EAL,[11,12] a multiplex monoclonal antibody-based approach[13] and physicochemical characterization.[14] The advantages of the last approach include increased sensitivity, accuracy, precision and linearity. In addition, the use of a non-immunologic approach allows the measurement of non-allergen components that may alter biological potency.

Previously, we reported the development of a multiplex assay for five GCr allergens (Bla g 1-Bla g 5) based on MRM technology.[14] The current work describes the development and optimization of a new MRM method for 10 additional GCr allergens, including isoforms of known allergens and vitellogenin. We applied this comprehensive MRM method to detect and quantify all GCr allergens in various background matrices including in commercial extracts and environmental samples. We also applied the method to evaluate the effect of extraction conditions and source materials on the heterogeneity of GCr extracts.

2 | METHODS

2.1 | Materials

Glycerinated German and American cockroach extracts were purchased from ALK-Abelló, Allermed Laboratories, Stallergenes-Greer and Jubilant HollisterStier (Table S1 in the Online Data Repository). These extracts presumably include all life stages of GCr, including nymphs, adult males, adult females and egg cases, as well as faeces and exuviae. Sequencing grade trypsin was obtained from Promega. Reference peptides (Table 1) were chemically synthesized in isotopically labelled forms (Thermo Fisher Scientific). All solvents were LC-MS grade (Fisher Scientific). Reagents for defatting and extraction were from Sigma.

2.2 | Source materials for extraction

Frozen whole-body GCrs were from Carolina Biological Supply Company (Burlington, NC), Stallergenes-Greer and North Carolina State University (NSCU). Gender-separated GCr faeces were from NSCU, and GCr egg cases were from Stallergenes-Greer. Faeces were obtained separately from adult males and adult females; they were not contaminated with shed cuticle from moulting nymphs (commonly referred to as "frass") or egg case remains. For analysis of environmental samples, settled dust was collected from homes of asthma patients in Baltimore using Mitest dust collectors (Indoor Biotechnologies), sieved and stored at −30°C until extraction.[15] All source materials were received on dry ice and stored at −80°C. Protein determinations were made by bicinchoninic acid (BCA) assay.

Multiple reaction monitoring method development was previously described.[14] Peptides were selected for each identified allergen on the basis of sequence uniqueness, ionizability and suitability for chromatography. Selected peptides were chemically synthesized as isotopically tagged peptides, where the carbon and nitrogen of N-terminal residues (R or K) were substituted for heavier isotopes (C-13 and N-15) so that fragments generated from isotopically tagged references can be distinguished from fragment ions of native counterparts. These labelled peptides were employed extensively for optimization of fragmentation parameters such as cone voltage, collision energy and dwell time, followed by selection of best representative fragment ions for quantification and qualification. The chromatographic separation was also optimized to achieve best-resolved peaks with the least interference.

Multiple reaction monitoring method was evaluated for precision, linearity and comparability to orthogonal methods. In addition, we used MRM to detect and quantify allergens in commercial extracts and extracts prepared from environmental samples. Because these extracts have components unsuitable for direct MRM analysis, buffer exchange was performed using 3 kDa MWCO spin filter (Millipore Sigma) by repeated washing with 50 mM ammonium bicarbonate (ABC). Allergen loss was estimated at 10%-15% (data not shown) using rBla g 2, rBla g 4 and rBla g 5 (Indoor Biotechnologies).

2.3 | Proteomic mapping using liquid chromatography-high-resolution mass spectrometry

We subjected 29 commercial extracts and 13 laboratory-made extracts to reduction by dithiothreitol (10 mM DTT; 60°C for 30 min),

TABLE 1 Reference peptides and MRM parameters

Allergen	Isoforms and variants	Peptide sequence	Retention time (min)	Cone voltage	Quantifier 1		Quantifier 2	
					[Precursor (M + H)++, Fragment ion (M + H)+]	Collison energy	[Precursor (M + H)++, Fragment ion (M + H)+]	Collison energy
Bla g 1	Bla g 1.0101	ALFNEK	9.81	30	[361.3, 390.5], [365.3, 398.5]	18	[361.3, 537.4], [365.3, 545.4]	12
		IIELIR	12.67	18	[378.7, 530.5], [383.7, 540.5]	12	[378.7, 643.6], [383.7, 653.6]	16
	Bla g 1.0201	SPEFQSIVETLK	14.76	13	[689.3, 589.3], [693.3, 597.3]	24	[689.3, 789.5], [693.3, 797.5]	24
Bla g 2	Bla g 2.0101	AIIVGPK	9.88	36	[349.2, 400.2], [353.2, 408.2]	14	[349.2, 513.4], [353.2, 521.4]	10
Bla g 3	UniProt ID: D0VNY6	EELTFDGVK	11.73	30	[519.1, 565.4], [523.1, 573.4]	14	[519.1, 666.4], [523.1, 674.4]	16
		FVEIDR	9.85	18	[389.5, 403.2], [394.5, 413.3]	14	[389.5, 532.3], [394.5, 542.3]	12
	Bla g 3.0101	LEDVPNVDIR	12.08	28	[585.3, 713.6], [590.3, 723.6]	18	[585.3, 927.4], [590.3, 937.4]	18
		NYFVEIDR	12.37	30	[528.3, 631.6], [533.3, 641.5]	16	[528.3, 778.7], [533.3, 788.7]	16
Bla g 4	Bla g 4.0101	AIEEDLK	9.07	20	[409.2, 615.5], [413.2, 623.5]	16	[409.2, 633.5], [413.2, 641.5]	12
Bla g 5	Bla g 5.0101	TPVLEIDGK	11.49	18	[486.3, 435.9], [490.3, 443.9]	18	[486.3, 674.5], [490.3, 682.5]	20
		VLGLPAIK	13.77	16	[405.7, 428.4], [409.7, 436.4]	18	[405.7, 598.6], [409.7, 606.3]	12
Bla g 6	Bla g 6.0101	EGNGYITTNVLR	12.02	44	[668.7, 703.4], [673.7, 713.4]	22	[668.7, 816.6], [673.7, 826.6]	22
		FLVEEDAEAMQQELR	14.98	26	[904.4, 804.5], [909.4, 814.6]	28	[904.4, 875.4], [909.4, 885.5]	26
	Bla g 6.0201	EGNGYITTAVLR	7.06	14	[647.2, 660.7], [652.2, 670.7]	22	[647.2, 773.6], [652.2, 783.6]	22
	Bla g 6.0301	SGSISTNMVEEILR	15.78	30	[768.3, 758.4], [773.3, 768.4]	26	[768.3, 1191.7], [773.3, 1201.7]	26
		TLEELIDEVDADK	12.41	30	[745.3, 791.4], [749.3, 799.4]	24	[745.3, 903.4], [749.3, 911.4]	24
Bla g 7	Bla g 7.0101	LAEASQAADESER		25	[688.8, 706.3], [692.8, 716.3]	20	[688.8, 777.4], [692.8, 787.4]	20
		IQLLEEDLER	13.28	38	[629.2, 903.5], [634.2, 915.3]	20	[629.2, 1016.6], [634.2, 1026.6]	20
		SLEVSEEK	8.37	18	[460.6, 591.34], [464.6, 599.4]	14	[460.6, 720.4], [464.6, 728.4]	14
Bla g 8	Bla g 8.0101	TFEDDGK	7.2	24	[406.1, 545.3], [410.1, 553.2]	16	[406.1, 563.3], [410.1, 571.3]	12
		EAFQLMDADK	12.35	16	[584.1, 575.3], [588.1, 583.3]	12	[584.1, 692.3], [588.1, 700.3]	12
		GSNVFSMFSQK	11.45	25	[616.1, 640.3], [620.1, 583.3]	14	[616.1, 727.4], [620.1, 735.4]	14
Bla g 9	Bla g 9.0101	VPFSHDDR	7.56	25	[486.7, 629.3], [491.7, 639.3]	25	[486.7, 776.4], [491.7, 786.4]	25
Bla g 11	Bla g 11.0101	TVTVGSDGK	7.05	30	[431.9, 463.3], [435.9, 471.3]	18	[431.9, 663.4], [435.9, 671.4]	10
		DIGDAFR	11.02	36	[396.9, 393.2], [401.9, 403.3]	18	[396.9, 565.3], [401.9, 575.3]	10
Vitellogenin		NVGDLSYSTSLVK	12.3	25	[691.7, 884.5], [695.7, 610.3]	20	[691.7, 1169.7], [695.7, 1177.7]	20

alkylation by iodoacetamide (50 mM IAA, 60°C for 30 min) and quenching (40 mM DTT; 60°C for 30 min). After an overnight trypsin digestion and trypsin inactivation, Rapigest was hydrolysed and the supernatant was concentrated in a vacuum concentrator. The digest was reconstituted in 0.1% formic acid and analysed by liquid chromatography-high-resolution mass spectrometry (LC/HRMS).

2.4 | LC/HRMS parameters

Extract digests (1-2 μg total protein) were loaded on a pre-column and separated on the UltiMate 3000RSLC nano system equipped with nano pump NCS-3000 and autosampler WPS-3000TPL (Thermo Scientific Dionex). The EASY-Spray analytical column (Acclaim PepMap: C18, 2 μm, 75 μm i.d., 75 cm long) was connected to the LC system with a Thermo Scientific Dionex nanoViper fingertight fitting. Column temperature was maintained at 40°C during all experiments. Injection, sample loading, column equilibration and autosampler wash conditions were kept consistent between the gradient durations and column lengths; flow rate during the gradient was 300 nL/min.

An Orbitrap Fusion Lumos MS was used for peptide MS/MS analysis. Survey scans of peptide precursors were performed at 350-1500 m/z at 120K FWHM resolution with a 4×10^5 ion count target and a maximum injection time of 50 ms The instrument was set to run with 3 s cycles for survey and MS/MS scans. After a survey scan, tandem MS was performed on the most abundant precursors with charge states from 2 to 6 and intensity > 5×10^3 by isolating them in the quadrupole. Collision-induced dissociation at constant 35% collision energy was used for fragmentation, and the resulting fragments were detected using the rapid scan rate in the ion trap. The automatic gain control target for MS/MS was set to 10^4 and the maximum injection time limited to 35 ms

Raw data were processed using Bionic software version 2.1.0.80, and MS/MS spectra were searched against *Blattella germanica* database (NCBI, 30 192 entry). The search rule included trypsin-specific peptides with up to two missed cleavages, and carbamidomethylation (+57.021 Da) of cysteine residues was set as a fixed modification. Oxidation of methionine residues (+15.9949 Da), acetylation of the protein N-terminus (+42.0106 Da) and deamidation of asparagine and glutamine (+0.984 Da) were treated as variable modifications. Precursor mass tolerance was 20 ppm, and product ions were searched at 0.5 Da tolerances.

2.5 | Preparation of allergen extracts

Frozen GCr bodies (nymphs and adult males and females) or egg cases were rinsed to remove debris and faeces, transferred to a disposable grinding chamber containing a small amount of dry ice and ground using an IKA Tube Mill control (twice, 30 seconds each, 25 000 rpm). Ground material was transferred to a 30-mL round-bottom glass tube and defatted at room temperature with at least 5 volumes of solvent (Table S2 in the Online Data Repository), for 8 cycles, 1 hour per cycle, on a shaking platform (300 rpm). Solvent was discarded, and defatted material was air-dried. Defatted powder was suspended in 20 volumes of extraction buffer (Table S2 in the Online Data Repository) in a 30-mL glass tube, agitated at 4°C for 72 hours on a shaker platform (750 rpm) and centrifuged at 800g for 5 minutes. The supernatant was transferred to 1.5-mL microcentrifuge tubes and spun at 18000g for 30 min. Supernatants were collected and stored at −20°C, either in the extraction buffer alone or with glycerol added to 50% (v/v).

GCr faeces (adult male or adult female) was pulverized in the presence of dry ice, as above, and defatted in at least 4 volumes of diethyl ether/ethyl acetate, 1:2 v/v, 4 times, 1.5 h each, at room temperature. Dried defatted powder was suspended in ABC at 4°C for 72 h with constant shaking. The extract was separated from the pellet by centrifugation. Extract was stored at −20°C, either in the extract buffer alone or with glycerol (50%, v/v) and phenol (0.4% w/v).

Environmental dust samples were extracted at 4°C (1:20, w/v, in PBS with 0.1% Tween20). After centrifugation, the supernatant was sterile-filtered and stored at −30°C.

For all MRM analyses, PBS was exchanged for ABC. Samples were subjected to overnight trypsinization (ratio 1:1); digestion was stopped by acidification and dried. Digests were reconstituted in 0.1% formic acid; the appropriate reference peptide cocktail was added, and LC/MRM-MS analysis was performed.

3 | RESULTS

3.1 | Multiple reaction monitoring mass (MRM) method development

3.1.1 | Isoform- or variant-specific allergen quantification

Most of the GCr allergens have sequence variants[16,17] but it is unknown whether these protein variants are represented in allergen extracts. Because our MRM method aimed to include all known allergens and their variants, we first analysed allergen extracts for protein composition and sequence diversity using high-resolution LC/MS (Orbitrap Fusion). From 42 allergenic extracts obtained from the US market or prepared by us, we identified 539 proteins and sequence variants (Table S3 in the Online Data Repository). Based on the proteomic map, most of the predicted allergen variants were undetectable in the analysed samples. Our analysis identified four Bla g 4 variants, three Bla g 2 variants and two variants each for Bla g 1 and Bla g 3. The variants of Bla g 1 and Bla g 3 were identifiable in all examined extracts, while the variants of Bla g 2 and Bla g 4 were detected in only a subset of samples. Bla g 6 was a minor component by mass in many of the analysed extracts, a finding that was confirmed using a LC-MRM/MS approach.

The allergens and variants for which the MRM method was developed are summarized in Table 1. The criteria for peptide selection

were previously described[14]; the main criterion is the uniqueness of the peptide sequence to the given variant. Those allergens or sequence variants confirmed to be frequently detectable were represented by at least one unique peptide. In total, 26 peptides representing all recognized GCr allergens (www.allergen.org) were identified and their isotopically tagged analogues were synthetically prepared for LC gradient and MRM method development and optimization.

The optimized MRM parameters are depicted in Table 1. We first established a calibration curve by spiking isotopically labelled analogue (SI) into digested extract at various concentrations. This approach differed from the one we used previously, which required synthetic native peptide (sNP) for analysis. We confirmed that the modified method is suitable and has comparable performance to our previous method.[14] Linearity was confirmed by spiking various concentrations of SI into GCr digests and plotting the MRM response against SI concentration. The MRM method was linear over the range 1-1000 fmol/μL with $R^2 > 0.97$ (representative data in Figure S1 in the Online Data Repository, panels A–C). Quantitation accuracy was confirmed by simultaneous quantification of recombinant allergens using the modified and unmodified MRM methods (data not shown). We also tested the effect of SI quantity on accuracy by spiking lower (10 fmol) and higher (100 fmol) concentrations of SI in GCr digests. Accuracy is robust within the linearity range (summarized in Figure S1 panels D, E, and F for Bla g 1.01, Bla g 2.01 and Bla g 5.01, respectively). Therefore, for this work, calibration was performed by a single point at 50 fmol/μL SI for each peptide. For Bla g 3 and Bla g 4, where the content of endogenous allergens is high, a fivefold dilution of GCr digest was prepared prior to spiking with SI.

3.1.2 | Evaluation of the LC-MRM/MS method

The methods evaluation for precision, accuracy and other pertinent parameters has been discussed.[14] Only additional evaluations are covered here.

Performance compared to orthogonal method

We compared Bla g 1 and Bla g 2 content measured by MRM to the content measured by commercially available ELISA (Indoor Biotechnologies) in parallel experiments on 20 different commercial GCr extracts. We found a strong correlation ($R > 0.9$) between LC/MRM-MS and ELISA, although there appears a systematic bias towards lower concentration of LC/MRM-MS over ELISA (Figure S2 in the Online Data Repository), as has been previously observed.[18]

Use of LC/MRM-MS method to detect and quantify allergens in different matrices

Commercial extracts: We examined 21 GCr allergen extract lots and 8 lots containing both GCr and American cockroach (Table S1 in the Online Data Repository). Expired extracts > 72 months were included to assess the performance of LC-MRM/MS on degraded samples. For the allergen content analysis, an equal total protein mass of each extract was run through the LC-MRM/MS protocol. Individual allergen content was determined, and the concentration in the original extract was calculated in ng/mL (Table 2). As expected, Bla g 1, Bla g 2 and Bla g 3, which are neither stage-specific nor sex-specific, are most abundant. Bla g 4, an adult male-specific allergen, also occurs in high abundance, suggesting that adult males were well-represented in these extracts. Bla g 3 and Bla g 4 account for 55%-85% of the allergen mass in these extracts, while Bla g 5 accounts for less than 0.5% of the content. Specific allergen content varies considerably among extracts from various sources (Table 2, and Figure S3 in the Online Data Repository). Finally, it appears that GCr allergens, as measured by this method, are stable under normal methods of storage (see, eg Table 2, Bla g 1 and Bla g 2 contents of manufacturer 2/lots 5, 6 and 7, among others). These MRM quantification results are largely confirmed by ELISA (data not shown).

Detection and measurement of allergens in environmental samples: LC-MRM/MS was optimized and verified to measure allergens in GCr aqueous extracts. We then determined the applicability of the method to measure allergens in environmental dust samples. Dust samples from 24 households (16 confirmed cockroach-infested and

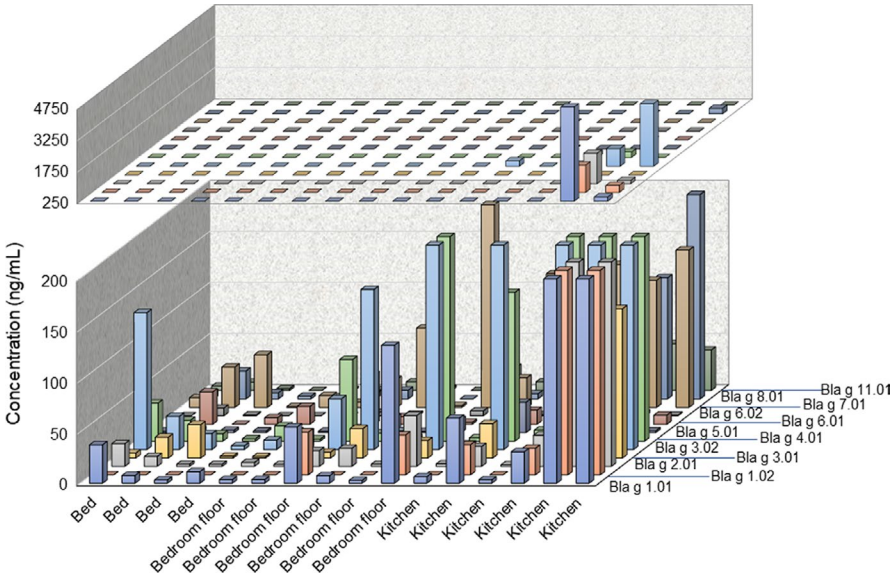

FIGURE 1 Application of the MRM method for detection and quantification of allergens in settled dusts of selected homes. Note different scale in top panel

TABLE 2 Quantification of detectable allergens and their isoforms in cockroach allergenic products obtained from four commercial sources (ng/mL)

Product #	Product type	Strength (w/v)	Age at MRM assay day	Bla g 1.0101	Bla g 1.0201	Bla g 2.0101	Bla g 3 UniProt ID: D0VNY6	Bla g 3.0101
	Manufacturer 1							
1	GCr	1:10	25	5380.1	643.3	2222.3	7680.7	43 603.5
2	GCr	1:10	17	4931.3	493.0	2062.2	6288.6	35 792.7
3	GCr	1:10	69	8670.0	1059.8	1504.1	4115.2	14 941.1
4	GCr	1:10	60	9310.7	1088.6	1475.1	4157.4	16 746.2
5	GCr	1:10	207	3949.8	473.5	2545.2	8702.1	47 787.0
6	Mix	1:10	25	3343.0	210.6	1106.7	4727.0	33 336.9
7	Mix	1:10	195	3269.2	356.3	757.3	5216.5	26 905.7
	Manufacturer 2							
1	GCr	1:20	33	6998.1	675.6	1347.9	4094.3	15 224.0
2	GCr	1:20	78	10 543.8	934.1	1147.0	6185.1	22 551.4
3	GCr	1:20	80	6216.0	601.4	1484.3	5973.7	35 896.3
4	GCr	1:20	80	6025.0	538.8	1457.3	4525.3	19 081.8
5	GCr	1:20	157	3055.5	274.3	1454.1	2259.8	7258.1
6	GCr	1:20	16	5832.6	524.2	1907.3	3536.2	10 947.2
7	GCr	1:20	16	6319.9	698.9	1950.6	3315.5	10 589.3
8	Mix	1:20	31	2598.7	176.3	493.1	2605.6	34 999.5
9	Mix	1:20	154	1677.8	183.4	509.0	1926.5	9646.5
	Manufacturer 3							
1	GCr	1:10	35	30 044.5	2548.1	14 211.2	27 218.7	199 862.8
2	GCr	1:10	93	25 893.4	2142.7	12 473.1	28 017.1	166 254.9
3	GCr	1:10	167	31 598.5	3153.7	7847.0	10 382.4	51 301.5
4	Mix	1:10	66	9556.5	543.8	3438.9	10 505.2	49 654.5
5	Mix	1:10	129	15 635.0	1782.6	5039.4	8667.2	60 580.7
	Manufacturer 4							
1	GCr	1:20	43	6933.9	307.6	5676.2	4603.1	23 518.7
2	GCr	1:20	63	5483.1	291.7	6105.1	4040.4	43 521.4
3	GCr	1:20	99	5096.9	205.1	4668.1	3839.7	52 029.7
4	GCr	1:20	150	2383.7	314.1	857.6	2386.0	40 008.0
5	GCr	1:20	27	5789.5	260.7	4490.6	5542.4	27 831.8
6	GCr	1:20	15	6181.3	296.5	5026.3	4040.8	16 100.7
7	Mix	1:20	42	3140.9	186.5	1736.5	1930.8	8687.5
8	Mix	1:20	150	750.7	96.8	492.3	1542.4	17 737.1

8 with no detectable Bla g 1 using ELISA) were obtained, extracted and subjected to LC-MRM/MS. Consistent with previous findings,[19-22] highest levels of GCr allergens are found in kitchens, but allergens are also detected in bed and bedroom samples (Figure 1). The mean allergen profile in the environmental samples is compared with the mean profiles from commercial extracts in Table 3. Relative quantities of Bla g 1, Bla g 2, Bla g 3, Bla g 4 and Bla g 11 are similar, while Bla g 5, Bla g 6, Bla g 7 and Bla g 8 are present in the environmental samples but largely absent for the commercial extracts.

3.1.3 | High-resolution LC/MS to explore level of heterogeneity in GCr extracts

We performed an in-depth proteomic evaluation using high-resolution LC/MS and identified 539 proteins from all commercial extracts included in this study (Table S3 in the Online Data Repository), but each extract has < 250 proteins identified. Of the total proteomic mass, 10%-15% is identifiable allergens (Bla g 1-Bla g 11). Several other proteins that bind IgE, such as aldolase, enolase, heat shock protein 70 kDa, triosephosphate isomerase and vitellogenin,[23] are

Bla g 4.0101	Bla g 5.0101	Bla g 6.0101	Bla g 6.0201	Bla g 6.0301	Bla g 7.0101	Bla g 8.0101	Bla g 9.0101	Bla g 11.0101	Vitellogenin
14 611.2	875.3	13.2	2.4	44.0	23.7	4.7	2.0	4621.3	336.7
14 937.7	959.7	13.0	0.9	505.4	22.8	1.4	2.2	3183.1	1080.5
18 901.4	601.7	4.3	0.8	467.2	5.3	3.2	2.5	2245.6	70.2
21 989.5	353.1	7.6	0.6	644.7	11.7	2.8	1.8	2302.0	116.9
16 808.3	178.5	4.7	1.8	813.9	11.5	3.9	1.6	4562.0	251.8
9465.5	350.2	3.6	1.4	124.1	16.4	99.9	0.9	3340.5	1955.0
5513.9	267.4	2.5	1.1	42.6	11.4	65.3	1.1	2567.7	21.0
8210.5	11.0	8.9	1.9	524.3	667.9	34.4	40 558.3	1060.8	61.5
4338.6	23.3	29.0	1.4	1150.9	292.6	33.1	0.9	2009.1	59.5
9294.4	47.5	18.0	1.9	7470.1	995.5	34.9	2.4	3917.9	119.8
8880.1	46.7	14.4	0.7	497.5	559.1	1.7	1.1	1720.7	160.6
9244.7	11.5	5.1	0.7	1904.9	507.1	1.2	0.2	993.8	160.7
11 603.0	36.7	11.0	1.1	443.6	879.9	4.2	0.6	1635.6	266.3
12 149.9	34.9	5.0	0.4	1543.0	377.0	2.0	1.1	1359.7	4289.4
5864.5	55.8	4.1	1.3	737.5	331.0	80.7	0.1	1269.3	26.4
7585.7	70.3	4.3	0.8	1314.2	18.4	52.3	7353.5	1337.7	184.3
62 473.5	695.4	12.5	0.8	158.7	27.4	3.5	2.8	33 589.4	4432.9
63 498.4	353.5	10.7	0.8	181.1	15.4	3.5	3.2	29 860.7	3466.6
26 031.6	51.2	7.2	0.5	49.5	14.0	5.3	8.2	22 448.1	784.4
9805.4	27.1	5.7	0.5	20.3	23.9	106.3	1.9	10 553.0	767.9
14 985.2	49.9	6.4	1.5	578.4	21.0	142.3	5.4	14 537.2	645.5
36 161.9	59.7	5.6	1.9	42.9	146.6	6.7	1.7	5506.9	3561.0
63 894.9	730.3	2.1	2.9	182.7	646.9	29.9	1.7	3638.2	2213.7
52 996.3	413.3	9.9	0.5	106.5	911.5	24.2	2.5	5120.3	1492.8
3699.3	6.8	2.3	0.4	69.9	1530.8	3.6	0.3	2283.2	312.6
49 934.2	286.5	5.8	1.8	54.8	60.6	12.5	1.3	3695.9	4539.2
53 557.5	123.5	7.7	0.5	68.7	55.2	4.9	2.1	2502.9	2146.0
11 923.8	46.9	2.6	1.7	49.6	12.3	491.0	0.5	1861.4	550.0
2547.4	110.1	0.9	0.6	46.6	20.0	51.9	0.2	1375.2	1138.1

abundant in commercial extracts. When these are considered, the proportion of overall IgE-binding protein in these extracts increases to 45%-50% of the total proteomic mass.

Relative protein quantification was performed on the commercial GCr extracts. This was achieved using a label-free approach based on normalized total ion intensity attributed to a given protein in an extract digest. An equimolar amount of synthetic bovine catalase was spiked to each extract digest to normalize variability during analysis. A heatmap was constructed using the 366 (out of 539 total) proteins most frequently detected in GCr extracts. Differences between extracts are readily apparent (Figure 2A). For example, products from manufacturer 1 have more proteins shown in the upper left quadrant of the heatmap (such as phosphoglycerate mutase, aminopeptidase, chemosensory protein, hexamerin, transferrin and various uncharacterized proteins), while manufacturer 4 has more proteins in the lower right quadrant (such as cathepsin D, lysosomal alpha-galactosidase, apolipophorin, superoxide dismutase and cytosol aminopeptidase). Gene ontology shows that most constituents of these extracts are enzymes or binding proteins (Figure 2B).

TABLE 3 Allergen profiles in GCr commercial extracts and in extracts prepared from environmental samples. Proportion is expressed as per cent of total allergenic mass detected. The per cent shown is the mean of determinations on several lots for each commercial extract and for samples from several homes for each environmental source, as indicated on the table

Allergen isoform/ variant	Manufacturer 1	Manufacturer 2	Manufacturer 3	Manufacturer 4	Bed	Bedroom	kitchen
N	7	9	5	4	4	6	6
Bla g 1.0101	6.9	13.1	8.0	6.2	14.0	9.1	13.2
Bla g 1.0201	0.8	1.3	0.7	0.3	ND	8.0	7.7
Bla g 2.0101	2.9	3.8	3.8	5.0	5.4	5.4	5.8
Blag 3 UniProt ID: D0VNY6	9.3	7.5	7.3	4.9	10.9	5.8	2.1
Bla g 3.0101	52.8	24.5	53.3	22.5	22.8	32.6	30.9
Bla g 4.0101	19.8	23.4	16.7	54.0	9.3	14.8	13.8
Bla g 5.0101	1.2	0.06	0.2	0.2	2.3	2.8	1.9
Bla g 6.0101	0.02	0.02	0	0.01	17.3	10.7	4.9
Bla g 6.0201	0	0	0	0	2.1	0.7	0.7
Bla g 7.0101	0.03	1.4	0.01	0.06	16.7	14.3	12.3
Bla g 8.0101	0	0.02	0	0.01	7.2	4.6	11.8
Bla g 11.0101	5.2	3.0	9.0	3.2	2.3	1.5	2.9

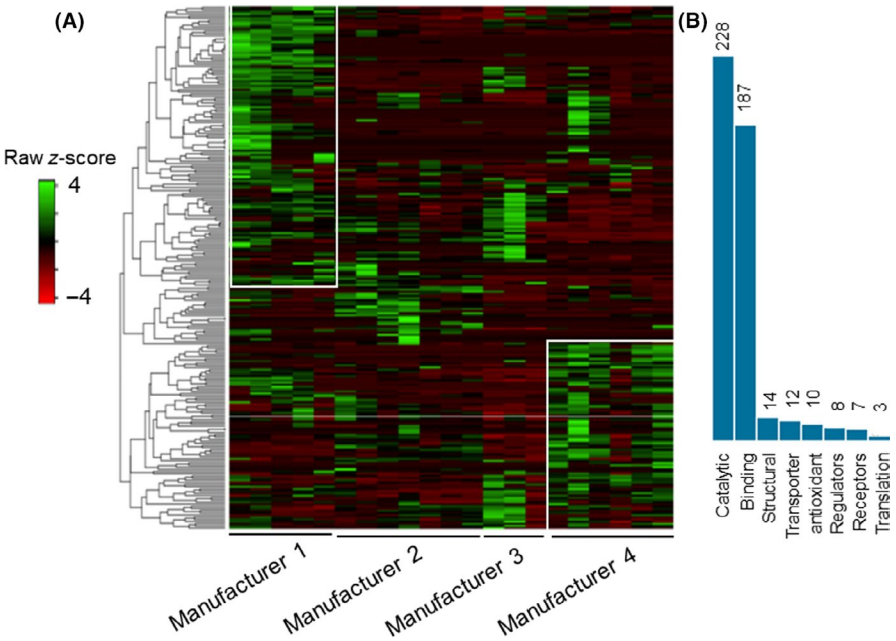

FIGURE 2 Proteomic map of typical GCr allergenic extracts. Panel A: Composition of various lots of GCr extracts from four commercial sources. Panel B: Functional ontology of the total proteome (n = 539 proteins) recovered from commercial extracts

3.1.4 | Compositional heterogeneity in GCr extracts

The effects of extraction conditions

Most, but not all, GCr source material extraction protocols begin with defatting steps.[23-29] We evaluated the impact of defatting and extraction choices on the extractable allergen quantity and profile. We subjected pulverized frozen GCr whole bodies to various combinations of 3 defatting solvents and 10 extraction buffers (Table S2 in the Online Data Repository). The total extractable proteins under these conditions are summarized in Table S4 in the Online Data Repository. An allergen-specific evaluation was performed using LC/MRM-MS, and the results are provided in Table E5. Further, Bla g 1 and Bla g 2 contents were assayed by specific ELISA for each of the extracts (Table S6 in the Online Data Repository). No single combination of extraction conditions was optimal across all of the tested parameters: total protein, MRM profile, specific allergen yield or completeness of extraction. Overall, differences in the resulting extracts were minor. Because of the yield and compatibility with downstream LC/MS analysis, we chose diethyl ether/ethyl acetate defatting solvent and 50 mM ABC buffer extraction for further assessment of the role of source materials in allergen extract composition.

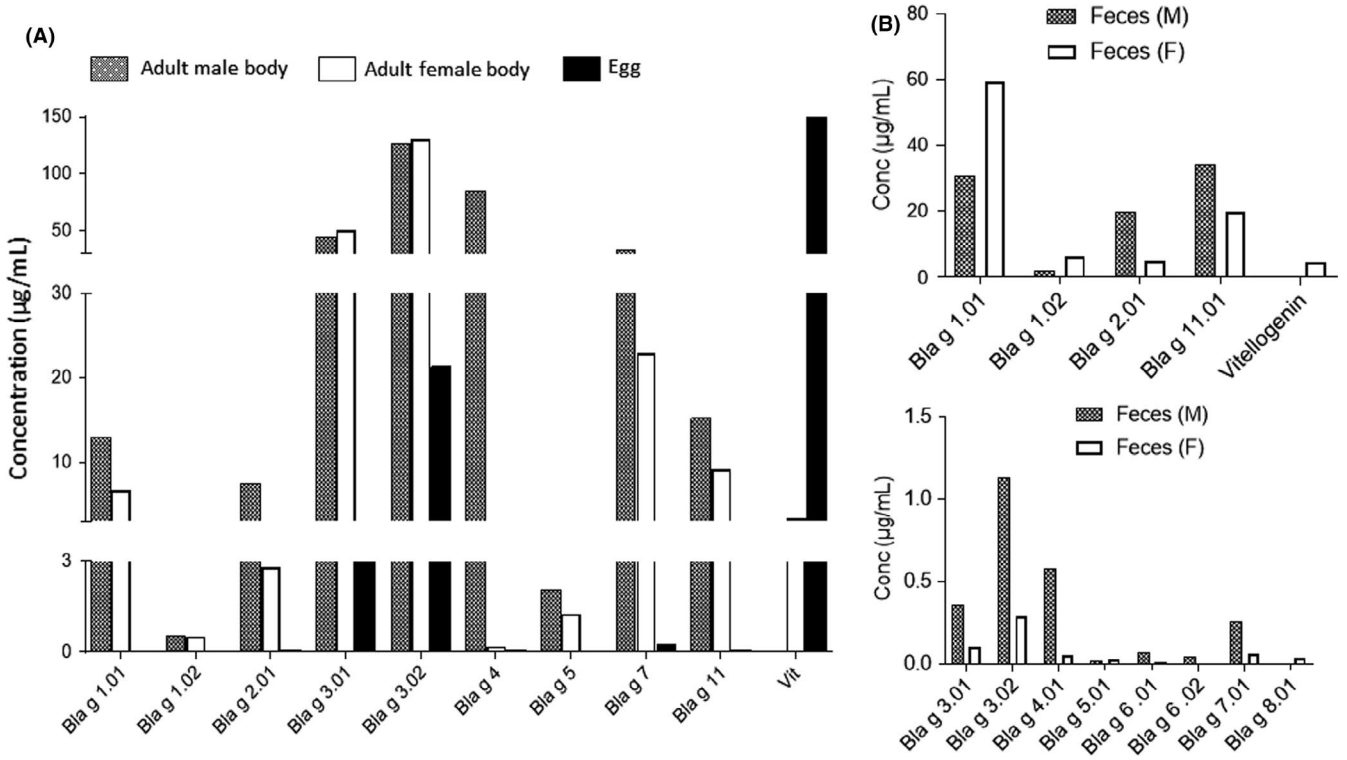

FIGURE 3 Impact of source materials used for extraction on the allergen content and profile of final product. Panel A compares the allergen profiles from whole-body extract based on gender of insects. The cockroaches were grown and treated the same way in the same location, and for comparison, the allergen content of female egg cases was included. Panel B: Comparison of excreted allergens from male and female German cockroach through faeces. Concentration is expressed in µg/mL. M and F: Adult male and adult female, respectively

The impact of source material used for extraction

We examined the effects of source materials on GCr extract allergen content using high-resolution LC/MS and LC/MRM-MS. For these analyses, we prepared extracts from whole bodies of GCr adults sorted by gender or from different providers; faeces from adult female and adult male GCr; and GCr eggs. While modest enrichment in male whole bodies over female was noted for several allergens, Bla g 4, a lipocalin previously localized to the male reproductive organs,[30] appears only in extracts from male GCr, and vitellogenin, an egg storage protein,[31] appears only in extracts from female GCr and eggs (Figure 3A). Faecal extracts had similar content differences, with Bla g 4 and vitellogenin found exclusively in faeces from adult male and female GCr, respectively (Figure 3B). Comparing the overall allergen levels, content in extracts from faeces was lower than levels in whole-body extracts, except for Bla g 1.01 and Bla g 11.

4 | DISCUSSION

Sensitization to GCr allergens causes asthma and rhinitis, which are significant public health burdens. Accurate skin test diagnosis and effective AIT for GCr allergy depend on the availability of well-characterized and consistent allergen extracts. The use of natural aqueous allergen extracts, as opposed to recombinant or synthetic allergens or allergen fragments, can be highly effective, safe and cost-effective,[32] but consistent manufacture will be enhanced by better methods of extract characterization. GCr extracts contain multiple allergens, and no one or two allergens are immunodominant.[6] In prior work, we have demonstrated that some GCr allergens can be measured using a multiplex monoclonal antibody-based immunoassay[13] and by MRM.[14] We now expand on our prior work by optimizing LC/MRM-MS to detect and quantify all known GCr allergens and their sequence variants.

Our previously described LC/MRM-MS protocol is expanded with minor modifications to be applicable to all known GCr allergens, and the performance of the modified assay continues to have excellent linearity, accuracy and precision.[14] Hence, properly designed and optimized, investigators can use LC/MRM-MS to detect and quantify allergens with precision and consistency because the identification and quantification are based on unique identifier sequences of the target allergens. Although the initial investment in analytical equipment and software is high, reagents are relatively inexpensive and easy to obtain, and development time is short compared to antibody-based assays. In our hands, LC/MRM-MS could be utilized on complex allergen mixes in a variety of matrices, without noticeable loss in performance characteristics.

One possible limitation of this approach—common to all physicochemical methods—is that it may lead to an overestimation of

allergen content, due to the detection of non-allergenic fragments and epitopes. When orthogonal testing is available, correlation with existing immunoassays is excellent, but ongoing validation will be necessary as this technique is applied to allergen extracts going forward.

Our work also indicates that LC/MRM-MS method can be complemented with high-resolution LC/MS to enhance product characterization beyond known allergens. Using the two techniques, we studied the proteomic compositions of various extracts from commercial sources as well as laboratory-made extracts and analysed the extent of heterogeneity among them, with special focus on factors that influence extract composition. Our analysis of three different defatting solvents and 10 different extraction protocols indicated that the methods are roughly comparable, although overall and allergen-specific yields varied. Source material selection had a much greater impact on extract composition than defatting and extraction protocols. Conventional whole-body GCr allergen source materials used by US manufacturers contain male and female insects, egg cases and faeces. Our analysis confirms[7] that the specific allergen content of a GCr allergen extract will depend on the relative proportions of these specific components. Extracts derived from source materials rich in faeces will contain more Bla g 1 and Bla g 11, whereas extracts from adult male whole bodies and adult female whole bodies/eggs will contain more Bla g 4 and vitellogenin, respectively. Furthermore, extracts derived from purified components can be expected to be much less diverse in their allergen content than extracts from crude whole bodies and excreta.

One further application of the LC/MRM-MS method in this study was the detection and evaluation of the GCr allergen profiles in selected homes of asthma patients around Baltimore. The profile of GCr allergens in the indoor environment is different from the distribution observed in extracts derived from mixed whole GCr bodies. Our study confirms that only some of the allergens in commercial GCr extracts reflect the GCr allergens to which urban asthma patients are exposed every day in their GCr-infested homes.

GCr infestations are difficult to control and especially difficult to eradicate in homes and schools,[15,33] and studies to improve GCr pest control are of critical public health importance. In addition, existing GCr allergen extracts are neither potent enough nor consistent enough to support their use as diagnostic and therapeutic agents for GCr-induced allergic disease[11]; aggressive research efforts to develop better GCr diagnostics and immunotherapeutics will be important to address this problem. As an analytical tool, LC/MRM-MS can be used to monitor and guide the results of both improved interventions in the environment and in allergen extract manufacture, towards improved data-guided solutions to GCr-induced asthma and allergic disease.

ACKNOWLEDGEMENTS

This work was supported by US Food and Drug Administration/Center for Biologics Evaluation and Research intramural funds. ND and SE were supported by an appointment to the Research Participation Program at the Center for Biologics Evaluation and Research administered by the Oak Ridge Institute for Science Education through an interagency agreement between the US Department of Energy and the US Food and Drug Administration.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Jay E. Slater https://orcid.org/0000-0001-7168-4276

REFERENCES

1. Bassirpour G, Zoratti E. Cockroach allergy and allergen-specific immunotherapy in asthma: potential and pitfalls. *Curr Opin Allergy Clin Immunol*. 2014;14(6):535-541.
2. Bridgewater J, Slater JE, Rabin RL. Regulation of allergen extracts in the United States. In Lockey RF, Ledford DK eds. *Allergens and Allergen Immunotherapy: Subcutaneous, Sublingual and Oral*. Boca Raton, FL: CRC Press;2020:354-360.
3. Rylander R, Holt PG. (1–>3)-beta-D-glucan and endotoxin modulate immune response to inhaled allergen. *Mediators Inflamm*. 1998;7(2):105-110.
4. Kale SL, Agrawal K, Gaur SN, Arora N. Cockroach protease allergen induces allergic airway inflammation via epithelial cell activation. *Sci Rep*. 2017;7:42341.
5. Eggleston PA. Cockroach allergy and urban asthma. *J Allergy Clin Immunol*. 2017;140(2):389-390.
6. Patterson ML, Slater JE. Characterization and comparison of commercially available German and American cockroach allergen extracts. *Clin Exp Allergy*. 2002;32(5):721-727.
7. Birrueta G, Frazier A, Pomes A, et al. Variability in German cockroach extract composition greatly impacts T cell potency in cockroach-allergic donors. *Front Immunol*. 2019;10:313.
8. Pomes A, Davies JM, Gadermaier G, et al. WHO/IUIS allergen nomenclature: providing a common language. *Mol Immunol*. 2018;100:3-13.
9. Glesner J, Filep S, Vailes LD, et al. Allergen content in German cockroach extracts and sensitization profiles to a new expanded set of cockroach allergens determine in vitro extract potency for IgE reactivity. *J Allergy Clin Immunol*. 2019;143(4):1474-1481.
10. FDA. Allergenic Products Advisory Committee meeting – 5 March 2001. https://wayback.archive-it.org/7993/20170403222320/https://www.fda.gov/ohrms/dockets/ac/cber01.htm#Allergenic%20Products. Published 2001. Accessed 6 May 2020.
11. Slater JE, James R, Pongracic JA, et al. Biological potency of German cockroach allergen extracts determined in an inner city population. *Clin Exp Allergy*. 2007;37(7):1033-1039.
12. James R, Mitchell H, Gergen PJ, Eggleston PA, Slater JE. Analyzing of ID50EAL data for the standardization of German cockroach allergen extracts in the U.S. *Arb Paul Ehrlich Inst Bundesamt Sera Impfstoffe Frankf A M*. 2006(95), 117-127; discussion 127, 155.
13. Khurana T, Dobrovolskaia E, Shartouny JR, Slater JE. Multiplex assay for protein profiling and potency measurement of German cockroach allergen extracts. *PLoS ONE*. 2015;10(10):e0140225.
14. Mindaye ST, Spiric J, David NA, Rabin RL, Slater JE. Accurate quantification of 5 German cockroach (GCr) allergens in complex extracts using multiple reaction monitoring mass spectrometry (MRM MS). *Clin Exp Allergy*. 2017;47(12):1661-1670.
15. Eggleston PA, Butz A, Rand C, et al. Home environmental intervention in inner-city asthma: a randomized controlled clinical trial. *Ann Allergy Asthma Immunol*. 2005;95(6):518-524.

16. Jeong KY, Lee H, Shin KH, et al. Sequence polymorphisms of major German cockroach allergens Bla g 1, Bla g 2, Bla g 4, and Bla g 5. *Int Arch Allergy Immunol*. 2008;145(1):1-8.
17. Jeong KY, Lee J, Lee IY, Ree HI, Hong CS, Yong TS. Analysis of amino acid sequence variations and immunoglobulin E-binding epitopes of German cockroach tropomyosin. *Clin Diagn Lab Immunol*. 2004;11(5):874-878.
18. Aebersold R, Burlingame AL, Bradshaw RA. Western blots versus selected reaction monitoring assays: time to turn the tables? *Mol Cell Proteomics*. 2013;12(9):2381-2382.
19. Arruda LK, Ferriani VP, Vailes LD, Pomes A, Chapman MD. Cockroach allergens: environmental distribution and relationship to disease. *Curr Allergy Asthma Rep*. 2001;1(5):466-473.
20. Chew GL, Perzanowski MS, Canfield SM, et al. Cockroach allergen levels and associations with cockroach-specific IgE. *J Allergy Clin Immunol*. 2008;121(1):240-245.
21. Chew GL. Assessment of environmental cockroach allergen exposure. *Curr Allergy Asthma Rep*. 2012;12(5):456-464.
22. Cohn RD, Arbes SJ Jr, Jaramillo R, Reid LH, Zeldin DC. National prevalence and exposure risk for cockroach allergen in U.S. households. *Environ Health Perspect*. 2006;114(4):522-526.
23. Chuang JG, Su SN, Chiang BL, Lee HJ, Chow LP. Proteome mining for novel IgE-binding proteins from the German cockroach (Blattella germanica) and allergen profiling of patients. *Proteomics*. 2010;10(21):3854-3867.
24. Jeong KY, Choi SY, Lee JH, et al. Preparation and characterization of an extract of german cockroach from a korean source. *Allergy Asthma Immunol Res*. 2013;5(2):102-105.
25. Yun YY, Ko SH, Park JW, Lee IY, Ree HI, Hong CS. Comparison of allergenic components between German cockroach whole body and fecal extracts. *Ann Allergy Asthma Immunol*. 2001;86(5):551-556.
26. Shin YS, Sohn JH, Kim JY, et al. Endotoxin is not essential for the development of cockroach induced allergic airway inflammation. *Yonsei Med J*. 2012;53(3):593-602.
27. Lee KE, Kim JW, Jeong KY, Kim KE, Yong TS, Sohn MH. Regulation of German cockroach extract-induced IL-8 expression in human airway epithelial cells. *Clin Exp Allergy*. 2007;37(9):1364-1373.
28. Stankus RP, O'Neil CE. Antigenic/allergenic characterization of American and German cockroach extracts. *J Allergy Clin Immunol*. 1988;81(3):563-570.
29. Lehrer SB, Horner WE, Menon P, Stankus RP. Comparison of cockroach allergenic activity in whole body and fecal extracts. *J Allergy Clin Immunol*. 1991;87(2):574-580.
30. Fan Y, Gore JC, Redding KO, Vailes LD, Chapman MD, Schal C. Tissue localization and regulation by juvenile hormone of human allergen Bla g 4 from the German cockroach, Blattella germanica (L.). *Insect Mol Biol*. 2005;14(1):45-53.
31. Tufail M, Takeda M. Molecular characteristics of insect vitellogenins. *J Insect Physiol*. 2008;54(12):1447-1458.
32. Cox L, Esch RE, Corbett M, Hankin C, Nelson M, Plunkett G. Allergen immunotherapy practice in the United States: guidelines, measures, and outcomes. *Ann Allergy Asthma Immunol*. 2011;107(4):pp. 289–299; quiz 300.
33. Sever ML, Arbes SJ Jr, Gore JC, et al. Cockroach allergen reduction by cockroach control alone in low-income urban homes: a randomized control trial. *J Allergy Clin Immunol*. 2007;120(4):849-855.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Mindaye ST, David NA, Zarkesh Esfahani SA, et al. Measurement of German cockroach allergens and their isoforms in allergen extracts with mass spectrometry. *Clin Exp Allergy*. 2020;50:741–751. https://doi.org/10.1111/cea.13604