Online supplement

ENVIRONMENTAL FUNGAL SENSITISATION ASSOCIATES WITH POORER CLINICAL OUTCOMES IN COPD

Pei Yee Tiew[1,2], Fanny Wai San Ko[3], Sze Lei Pang[4], Sri Anusha Matta[4], Yang Yie Sio[4], Mau Ern Poh[5], Kenny J. X. Lau[6], Micheál Mac Aogáin[1], Tavleen Kaur Jaggi[1], Fransiskus Xaverius Ivan[1], Nicolas E. Gaultier[6], Akira Uchida[6], Daniela I. Drautz-Moses[6], Huiying Xu[7], Mariko Siyue Koh[2], David Shu Cheong Hui[3], Augustine Tee[8], John Arputhan Abisheganaden[7], Stephan C. Schuster[6], Fook Tim Chew[4] and Sanjay H. Chotirmall[1#].

SUPPLEMENTARY MATERIALS AND METHODS

Ethics approval: Reference numbers pertaining to ethical approvals at each site was as follows: CIRB 2016/2549 (2013/184/C), CIRB 2016/2715, CIRB 2017/2933, CIRB 2017/2109 (all mutually recognized by DSRB, Singapore), UMMC 2018725-6524 (Malaysia), CREC 2011.146, CREC 2015.164 and CREC 2018.042 (Hong Kong). Non-diseased (healthy) control recruitment was approved by Nanyang Technological University (NTU) and UTAR Institutional Review Boards (IRBs) under IRB-2017-12-010 (Singapore), U/SERC/03/2015 and U/SERC/03/2016 (Malaysia).

Venous blood sampling and processing: Venous blood was collected at recruitment. Blood was centrifuged at 1300g for 10 minutes at 18°C, plasma was isolated and stored in aliquots at -80°C until further processing as described.

Total Immunoglobulin (Ig)-E: Total IgE was measured using the Human IgE ELISA Kit (Abcam ab108650) as per manufacturers instructions. Briefly, twenty microlitres of plasma from each patient (in duplicate) were screened against a set of standards within each microplate.

Crude fungal allergen preparation: Fungi were homogenised with phosphate-buffered saline (PBS) using mortar and pestle. The homogenised fungi were incubated at 4°C for 4h followed by centrifugation to isolate the supernatant. Glycerol (20%) was added to supernatants and filtered. Protein concentrations were determined by Bradford assays and stored at -20°C for immunological assessment.

Immuno-dot blot assay for specific-IgE (sIgE) measurements: Immuno-dot blot assays were used to assess for specific-IgE (sIgE) responses to the crude proteins *Dermatophagoides farinae* (Der f), *Dermatophagoides pteronyssinus* (Der p), *Blomia tropicalis* (Blo t), *Elaeis guineensis*, Panicoids,

Pooids, Chloroids, Weeds, *Blattella germanica* (Bla g), *Periplaneta Americana*, *Curvularia*, *Penicillium*, *Aspergillus fumigatus* (*A..fumigatus*) including air fungi and *Aspergillus* recombinant proteins as per published methodology by our group and others [1-4]. The recombinant proteins used in this study included: Asp f 1 (M83781), Asp f 2 (U56938), Asp f 4 (AJ001732), Asp f 6 (U53561), Asp f 8 (AJ224333), Asp f 15 (AJ002026) and Asp f 17 (AJ224865). Briefly, each allergen was blotted onto nitrocellulose membranes (in duplicate at a standardized concentration of 0.25mg/allergen) with Bovine Serum Albumin (BSA) used as protein control and PBS included as a negative control. Membranes were blocked with 1X PBS 0.1% Tween-20 for 40 minutes and incubated overnight (16 hours) with diluted plasma (1:8 in PBS) at 4°C. Washing steps were performed with 1X PBS 0.05% Tween-20 for 15 minutes followed by 7-minute washes in replicate. Subsequently, membranes were incubated with anti-human IgE antibodies (1:1000) conjugated with alkaline phosphatase (Sigma Aldrich, USA) for 2h. Nitroblue tetrazolium (NBT)/5-bromo-4-chloro-3'-indolyphosphate (BCIP) solution (Thermo Fisher Scientific) was then added for detection of alkaline phosphatase activity for 10 minutes and data analyzed with Syngene imaging software. Inter and intra assay reproducibility was 90% and 95% respectively and specific Ig-E binding is presented as optical density (OD) intensity.

Isolation and confirmatory identification of outdoor air fungi

Top outdoor air fungi: Based on outdoor air metagenomic profiles as described by Gusareva *et al*, we cultivated the top 8 fungi for further immunological analysis[5]. *Schizophyllum commune* was collected from the National Institute of Education Carpark 3 at Nanyang Technological University, *Trametes sanguinea* from Kent Ridge Park, Singapore. Permission for fungal specimen collection was filed under NParks Permit NP/RP19-078. *Aspergillus spp.* (*A. terreus*, *A. sydowii*), *Cladosporium spp.* (*C. tenussimum*, *C. xantochromaticum*), *Fusarium spp.* (*F. solani*, *F. moniliforme*), *Neurospora spp.*, *Byssochlamys spectabilis* and *Penicillium spp.* were directly isolated from air samples collected at a rooftop balcony at Nanyang Technological University, Singapore using a Spin-Air agar impactor (IUL S.A., Barcelona, Spain) at a sampling height of 156cm. Air was impacted onto various nutrient media including Malt Extract Agar (MEA) (Sigma-Aldrich, USA), cellulose agar (Sigma-Aldrich, USA),

Potato Dextrose Agar (Sigma-Aldrich, USA), Reasoner's 2A Agar (Becton-Dickinson, USA) and Tryptic Soy Agar (Becton-Dickinson, USA) in 90mm Petri dishes at an air flow rate of 100L/min for 1 min following protocols described by Sánchez-Muñoz, Muñoz-Vicente [6]. Impacted agar plates were then incubated at 25°C for 3-5 days and single fungal colonies selected and re-streaked onto fresh MEA plates. The selection and streaking process was repeated for 3 generations until plate morphology appeared to represent a single fungal isolate and to ensure the isolate species was free from contamination. Each of the fungal species isolated was then confirmed using molecular analysis as described.

*Fungal species identification***:** DNA was extracted from single fungal isolates using the Qiagen PowerWater DNA extraction kit (Qiagen, Germany) and subsequently amplified by polymerase chain reaction (PCR) targeting the ITS-4 and ITS-5 (internal transcribed spacer) rRNA genes. ITS-4 (5'-TCCTCCGCTTATTGATATGC-3') and ITS-5 (5'-GGAAGTAAAAGTCGTAACAAGG-3') primers were used and PCR conditions were as follows: 2x KAPA HiFi Hotstart Ready Mix (12.5 µl) (Kapa Biosystems, USA), primers (for both forward and reverse primers each 0.75µl of 10µM), template DNA (1µl of 5 ng/µl) and H_2O (10 µl) for each reaction (total 25 µl). PCR cycles for ITS-4 and ITS-5 were performed as follows: 95°C for 3 minutes, 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds with a final elongation step at 72ºC for 5 minutes. The PCR products were sent to AITbiotech Singapore for Sanger sequencing and read quality analysed with 4Peaks (Mekentosj, Amsterdam) and Chromas (Technelysium, Australia) where sites below Phred score 20 were filtered and trimmed. Forward and reverse reads were merged using BioEdit v7.2.6.1 [7] and blastn [8] used to query the sequences to confirm species identity. Only once each species was confirmed was the fungal extract used for further immunological analyses.

Indoor Allergen mapping: Data for 977 established allergens were extracted from the Allergen Nomenclature (http://www.allergen.org/) using the R "rvest" package [9]. Nucleotide sequences of the 786 allergens whose GenBank Accession was available were retrieved using the R "ape" package [10]. The raw paired-end Illumina HiSeq sequencing reads with a minimum Phred score of 20 and >30bp length were selected and adapter-trimmed using Cutadapt (version 1.8.1) [11]. The trimmed reads were then mapped against the GRCh38 human reference genome using bowtie2 (version 2.3.3.1) [12]. The unmapped non-human reads were extracted using samtools (version 1.3.1) and bedtools2 (version 2.8.0), and, aligned to the 786 allergens nucleotide sequences using bowtie2 in its local mode [13, 14]. Samtools was next used to calculate the number of alignments with mapping quality >30 (raw hits). The number of raw hits (H) to a specific allergen's nucleotide sequence was then normalized using similar methodology to ShortBRED [15], whereby read length (R), allergen nucleotide sequence length (L), and sequencing depth were adjusted after removal of human reads (N) according to the following formula:

$$C = \frac{H}{\left(\frac{|L-R|}{10^3}\right)\left(\frac{N}{10^6}\right)} = \frac{H}{|L-R| \cdot N} \times 10^9$$

The normalized count (relative abundance) is then presented in units of reads per kilobase of reference sequence per million sample reads (RKPM).

REFERENCES

1. Kidon MI, Chiang WC, Liew WK, Ong TC, Tiong YS, Wong KN, Angus AC, Ong ST, Gao YF, Reginald K, Bi XZ, Shang HS, Chew FT. Mite component-specific IgE repertoire and phenotypes of allergic disease in childhood: the tropical perspective. *Pediatr Allergy Immunol* 2011: 22(2): 202-210.

2. Chan SL, Ong TC, Gao YF, Tiong YS, Wang de Y, Chew FT, Mok YK. Nuclear magnetic resonance structure and IgE epitopes of Blo t 5, a major dust mite allergen. *J Immunol* 2008: 181(4): 2586-2596.

3. Gao YF, Wang de Y, Ong TC, Tay SL, Yap KH, Chew FT. Identification and characterization of a novel allergen from Blomia tropicalis: Blo t 21. *J Allergy Clin Immunol* 2007: 120(1): 105-112.

4. Batard T, Hrabina A, Bi XZ, Chabre H, Lemoine P, Couret MN, Faccenda D, Villet B, Harzic P, Andre F, Goh SY, Andre C, Chew FT, Moingeon P. Production and proteomic characterization of pharmaceutical-grade Dermatophagoides pteronyssinus and Dermatophagoides farinae extracts for allergy vaccines. *Int Arch Allergy Immunol* 2006: 140(4): 295-305.

5. Gusareva ES, Acerbi E, Lau KJX, Luhung I, Premkrishnan BNV, Kolundzija S, Purbojati RW, Wong A, Houghton JNI, Miller D, Gaultier NE, Heinle CE, Clare ME, Vettath VK, Kee C, Lim SBY, Chenard C, Phung WJ, Kushwaha KK, Nee AP, Putra A, Panicker D, Yanqing K, Hwee YZ, Lohar SR, Kuwata M, Kim HL, Yang L, Uchida A, Drautz-Moses DI, Junqueira ACM, Schuster SC. Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proc Natl Acad Sci U S A* 2019: 116(46): 23299-23308.

6. Sánchez-Muñoz M, Muñoz-Vicente M, Cobas G, Portela R, Amils R, Sánchez B. Comparison of three high-flow single-stage impaction-based air samplers for bacteria quantification: DUO SAS SUPER 360, SAMPL'AIR and SPIN AIR. *Analytical Methods* 2012: 4.

7. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 1999: 41: 95-98.

8. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* 1997: 25(17): 3389-3402.

9. Pomes A, Davies JM, Gadermaier G, Hilger C, Holzhauser T, Lidholm J, Lopata AL, Mueller GA, Nandy A, Radauer C, Chan SK, Jappe U, Kleine-Tebbe J, Thomas WR, Chapman MD, van Hage M, van Ree R, Vieths S, Raulf M, Goodman RE, Sub-Committee WIAN. WHO/IUIS Allergen Nomenclature: Providing a common language. *Molecular immunology* 2018: 100: 3-13.

10. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 2019: 35(3): 526-528.

11. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. . *EMBnet Journal* 2011: 17: 10-12.

12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012: 9(4): 357-359.

13. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009: 25(16): 2078-2079.

14. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010: 26(6): 841-842.

15. Kaminski J, Gibson MK, Franzosa EA, Segata N, Dantas G, Huttenhower C. High-Specificity Targeted Functional Profiling in Microbial Communities with ShortBRED. *PLoS computational biology* 2015: 11(12): e1004557.

Table E1 Demographic table of the COPD patient cohort participating in the indoor home air and surface sampling (n=11). Data are presented as number of homes/patients (n) (with percentage; %) or mean ± standard deviation (SD) or median (interquartile range; IQR). BMI: body mass index; FEV_1: forced expiratory volume in the 1st second; FVC: forced vital capacity.

Demographic	
Number of patient homes sampled	11
Age (years), Mean ±SD	69 ± 3
Sex (Male), n (%)	11 (100)
BMI (kg/m^2), Mean ±SD	22.5 ±3.7
Smoking status, n (%)	
Current	5 (45.5)
Ex-smoker	6 (54.5)
Never	0 (0.0)
Smoking pack years, Median (IQR)	50.0 (47.5-65.0)
FEV_1 % predicted, Mean ±SD	57.7±21.2
FEV_1/FVC (% predicted), Mean ±SD	56.7 ±12.5
Hospitalization in the year preceding study, n (%)	
Yes	3 (27)
No	8 (73)
Exacerbation number in the year preceding study, n (%)	
0-1	10 (91)
>1	1 (9)
COPD assessment test (CAT), Mean ±SD	12.3±11.1
Blood Eosinophil ($x10^9/L$)	0.21 (0.09-0.42)
Total IgE (IU/ml)	26.2 (5.0-251.6)
Treatment, n (%)	
LAMA	4 (36)
LAMA/LABA	4 (36)
LAMA/LABA/ICS	3 (28)

Table E2 Table showing systemic specific IgE (sIgE) binding (expressed as OD intensity) for each tested allergen within the three defined patient clusters. Data are presented as median (interquartile range; IQR) with p-value indicated where significant. OD: optical density.

Allergens	**Total, OD,**	**HS fungal predominant, OD**	**LS, OD**	**MS *Blomia* predominant, OD**	**p-value**
	Median (IQR)	**Median (IQR)**	**Median (IQR)**	**Median (IQR)**	
Elaeis guineensis	2.59 (1.55-6.51)	**3.31 (1.68-9.44)**	2.55 (1.55-7.5)	2.47 (1.48-5.55)	**0.01**
Panicoids	5.20 (2.01-16.92)	7.98 (2.49-22.24)	4.60 (1.86-16.57)	4.76 (1.95-15.52)	ns
Pooids	5.56 (2.35-17.94)	7.49 (3.02-23.24)	4.93 (2.16-15.88)	5.22 (2.26-15.86)	ns
Chloroids	4.09 (1.85-12.78)	5.22 (2.1-18.11)	3.40 (1.66-10.47)	3.75 (1.84-11.77)	ns
Weeds	2.51 (1.48-6.06)	**3.39 (1.77-6.49)**	1.93 (1.31-6.15)	2.35 (1.5-5.66)	**0.04**
Blomia tropicalis	4.88 (2.25-15.6)	5.34 (2.21-16.8)	3.68 (1.86-9.55)	**5.89 (2.67-17.52)**	**0.05**
Dermatophagoides farina	4.34 (1.96-12.71)	4.56 (2.27-12.9)	4.27 (1.65-11.79)	4.17 (1.89-13.32)	ns
Dermatophagoides pteronyssinus	4.20 (1.92-9.53)	4.92 (2.21-11.32)	3.43 (1.66-8.56)	4.23 (1.84-9.62)	ns
Blattella germanica	2.63 (1.57-6.06)	2.91 (1.87-6.93)	2.31 (1.4-6.06)	2.62 (1.57-5.65)	ns
Periplaneta americana	3.37 (1.65-7.56)	3.83 (2.03-8.59)	3.29 (1.48-6.18)	3.24 (1.59-8.44)	ns
Curvularia spp.	2.35 (1.43-5.13)	2.68 (1.63-5.44)	2.05 (1.26-5.77)	2.32 (1.45-4.65)	ns
Penicillium spp.	2.58 (1.48-5.98)	**3.19 (1.82-6.92)**	2.30 (1.33-6.16)	2.51 (1.45-5.53)	**0.05**
Aspergillus fumigatus	2.34 (1.44-5.26)	**3.39 (1.64-6.38)**	2.04 (1.42-5.05)	2.31 (1.44-4.75)	**0.02**
rAsp f 1	2.66 (1.26-7.51)	3.23 (1.15-8.98)	2.90 (1.47-7.81)	2.31 (1.14-6.78)	ns
rAsp f 2	0.13 (0.02-1.02)	0.19 (0.02-1.69)	0.09 (0.03-0.93)	0.12 (0.02-0.86)	ns
rAsp f 6	2.40 (1.48-6.26)	2.87 (1.54-7.33)	2.11 (1.49-6.18)	2.34 (1.4-5.81)	ns
rAsp f 8	2.41 (1.42-6.03)	**3.00 (1.58-7.3)**	2.29 (1.39-5.83)	2.25 (1.31-5.69)	**0.03**
rAsp f 15	2.83 (1.48-6.63)	3.50 (1.59-7.81)	2.56 (1.49-6.04)	2.57 (1.43-6.31)	ns
rAsp f 17	2.44 (1.41-5.71)	**3.10 (1.51-6.88)**	2.11 (1.32-4.82)	2.39 (1.43-5.08)	**0.04**

Figure E1: Overall summary of the allergen panels used in this study. HDM: house-dust mite, Der f: *Dermatophagoides farinae*, Der p: *Dermatophagoides pteronyssinus*, Blo t: *Blomia tropicalis*, Bla g: *Blattella germanica*, and rAsp f: recombinant *Aspergillus fumigatus*.

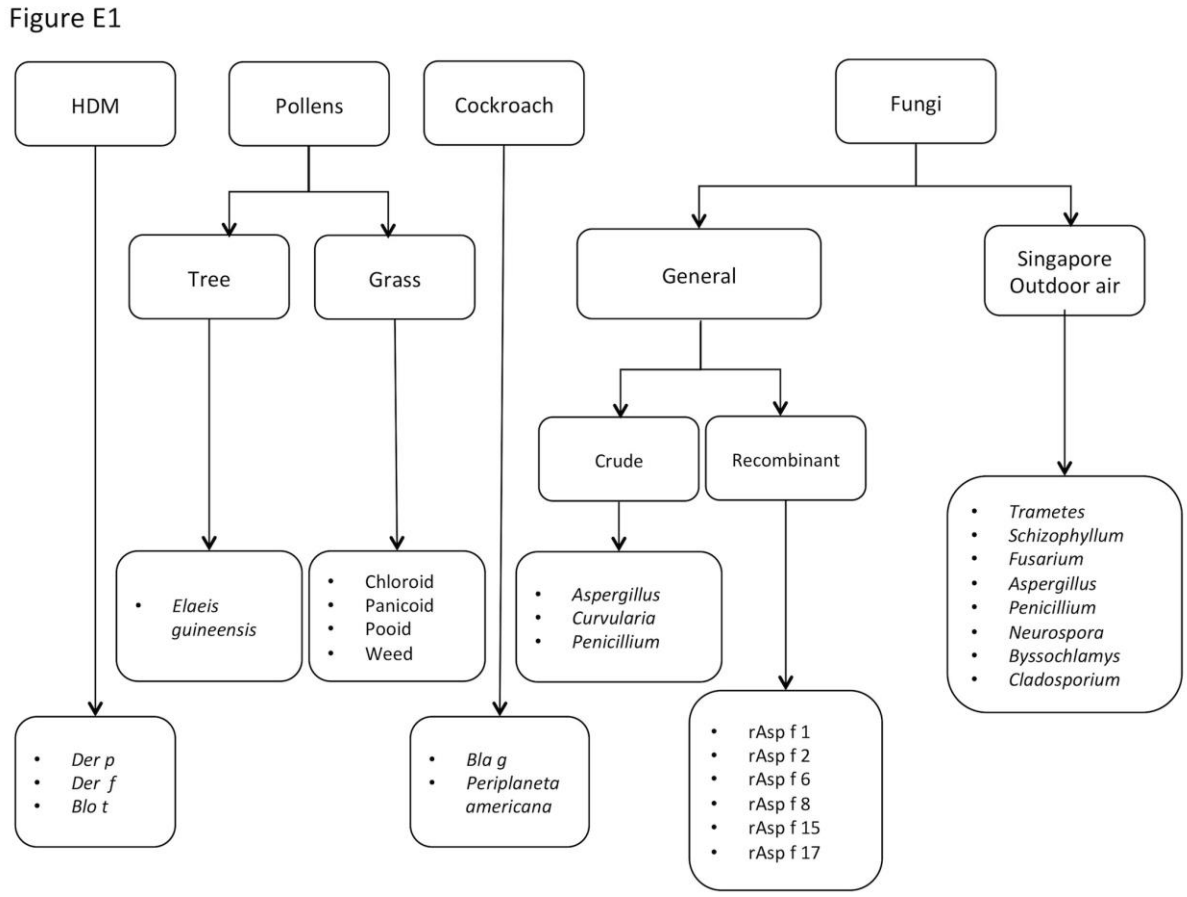


Figure E2: Sensitization rates to (a) pollens, (b) house dust mite and cockroach allergens (except *Blattella germanica*) does not differ based on frequent exacerbator status in COPD. Systemic specific IgE (sIgE) binding is expressed as $\log_{10}$ OD intensity. Boxplot colouration corresponds to exacerbator status: non-frequent exacerbator (grey) and frequent exacerbator (pink). HDM: house-dust mite, ns: non-significant. *p<0.05.

Figure E2

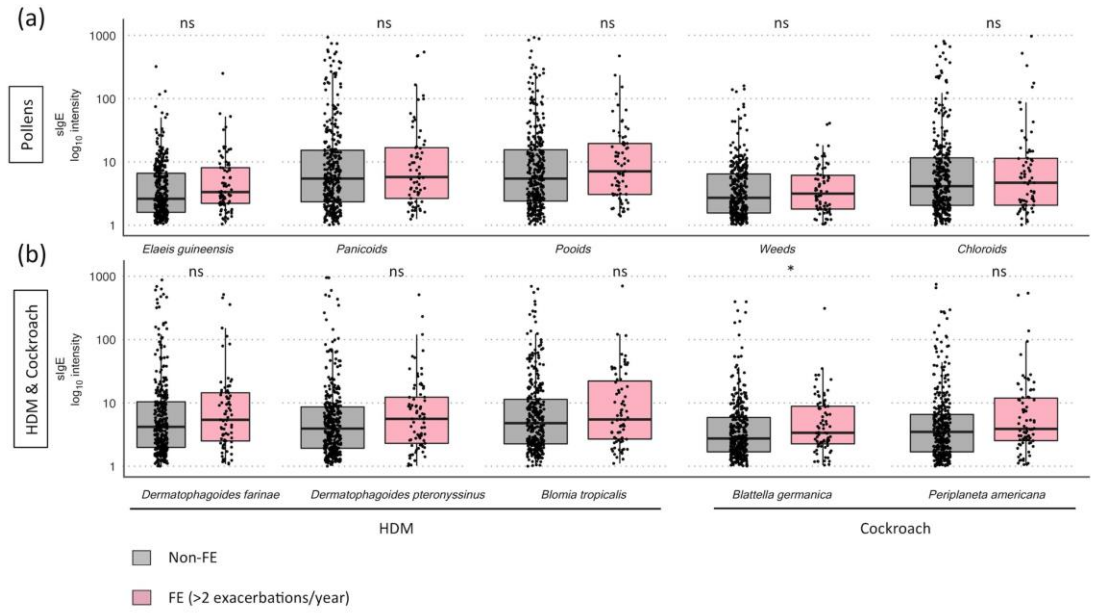
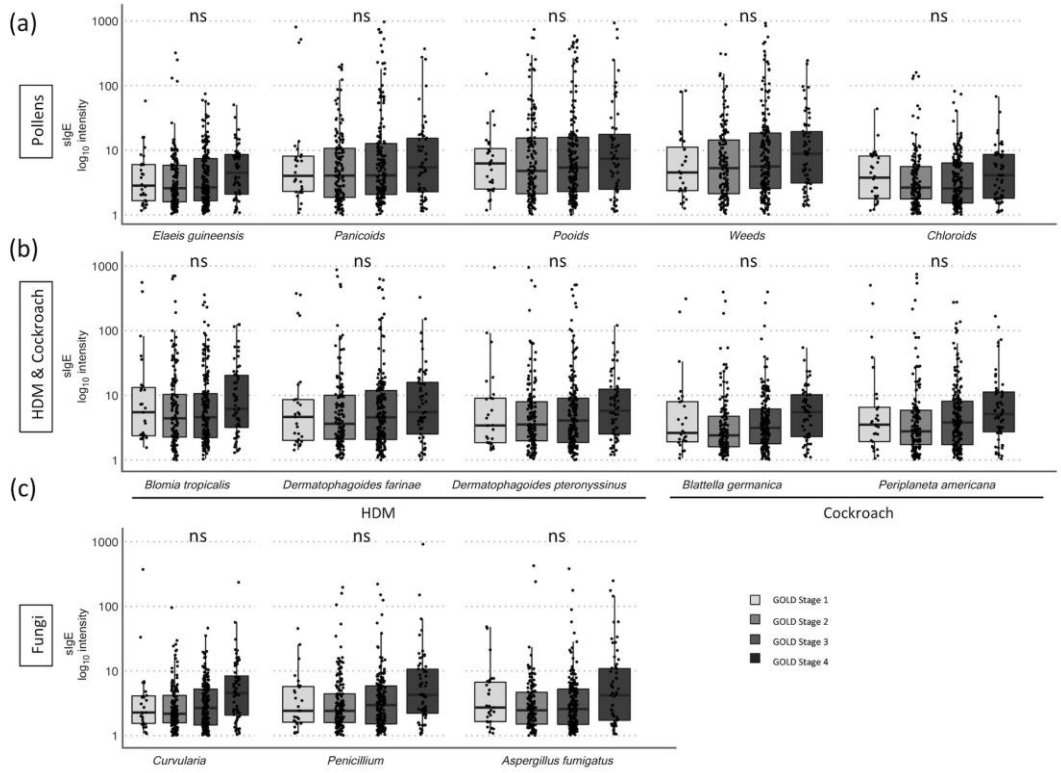


Figure E3: Sensitization rates to (a) pollens, (b) house dust mite, cockroach and (c-e) fungal allergens (crude and recombinant) does not differ based on GOLD (lung function) stage. Systemic specific IgE (sIgE) binding is expressed as log_{10} OD intensity. Boxplot colouration corresponds to GOLD (lung function; FEV_1) stage: GOLD stage 1 (light grey), GOLD stage 2 (medium grey), GOLD stage 3 (grey), GOLD stage 4 (dark grey). ns: non-significant.

Figure E3



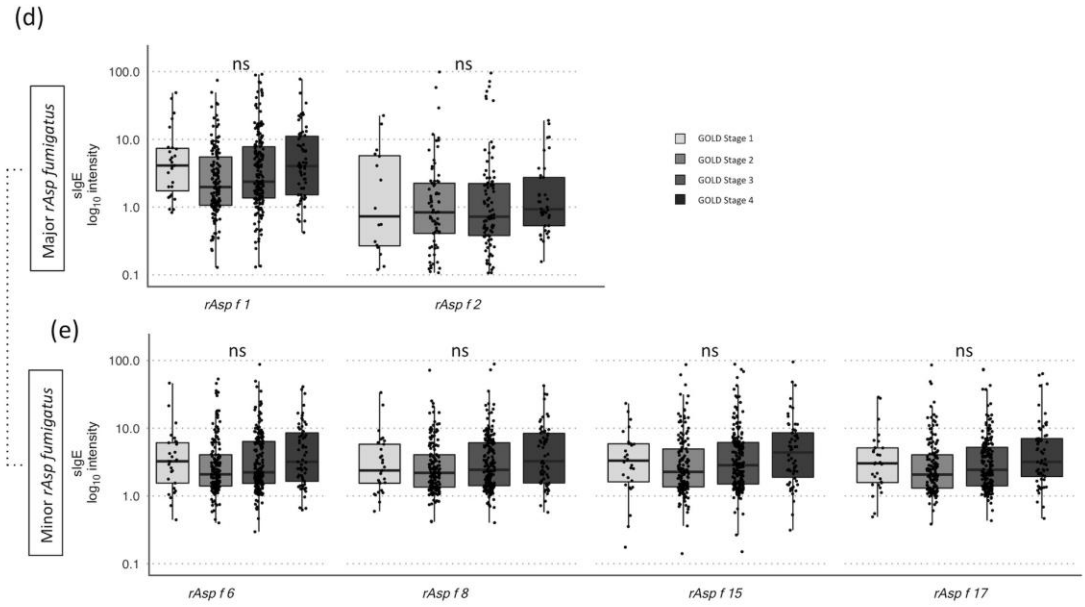
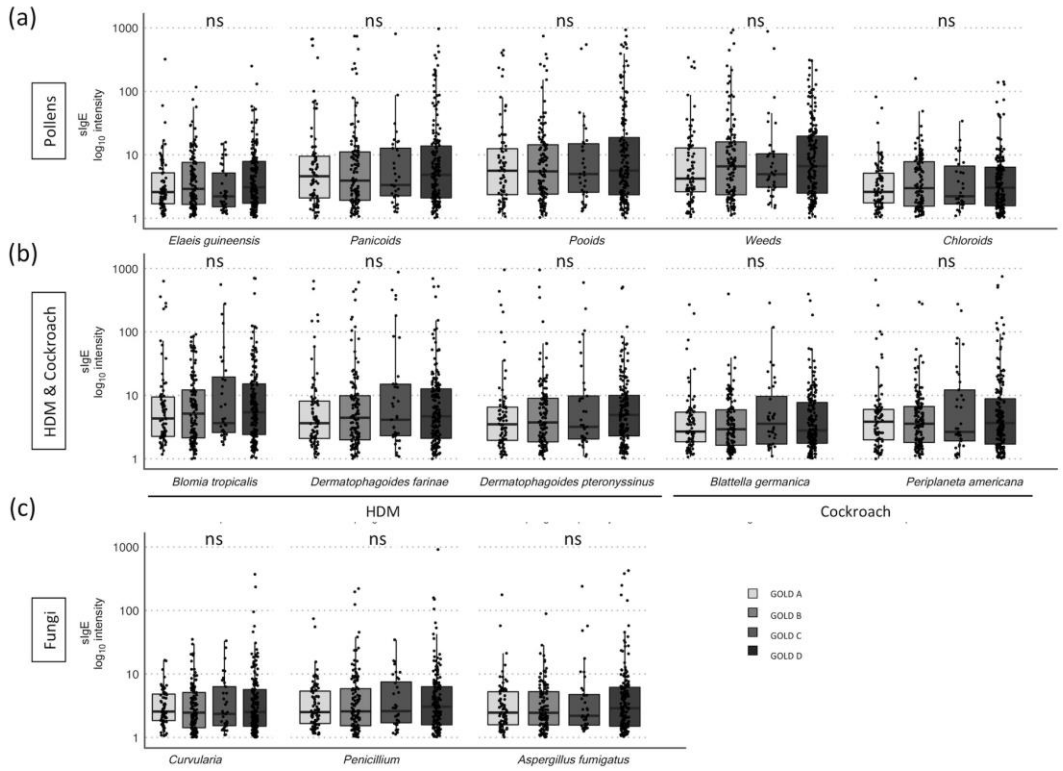

(d)
Major rAsp fumigatus
sIgE
log10 intensity
100.0
10.0
1.0
0.1
ns
ns
rAsp f 1
rAsp f 2
GOLD Stage 1
GOLD Stage 2
GOLD Stage 3
GOLD Stage 4
(e)
Minor rAsp fumigatus
sIgE
log10 intensity
100.0
10.0
1.0
0.1
ns
ns
ns
ns
rAsp f 6
rAsp f 8
rAsp f 15
rAsp f 17

Figure E4: Sensitization rates to (a) pollens, (b) house dust mite, cockroach and (c-e) fungal allergens (crude and recombinant) does not differ based on GOLD group (ABCD). Systemic specific IgE (sIgE) binding is expressed as $\log_{10}$ OD intensity. Boxplot colouration corresponds to GOLD group: GOLD A (light grey), GOLD B (medium grey), GOLD C (grey), GOLD D (dark grey). ns: non-significant.

Figure E4



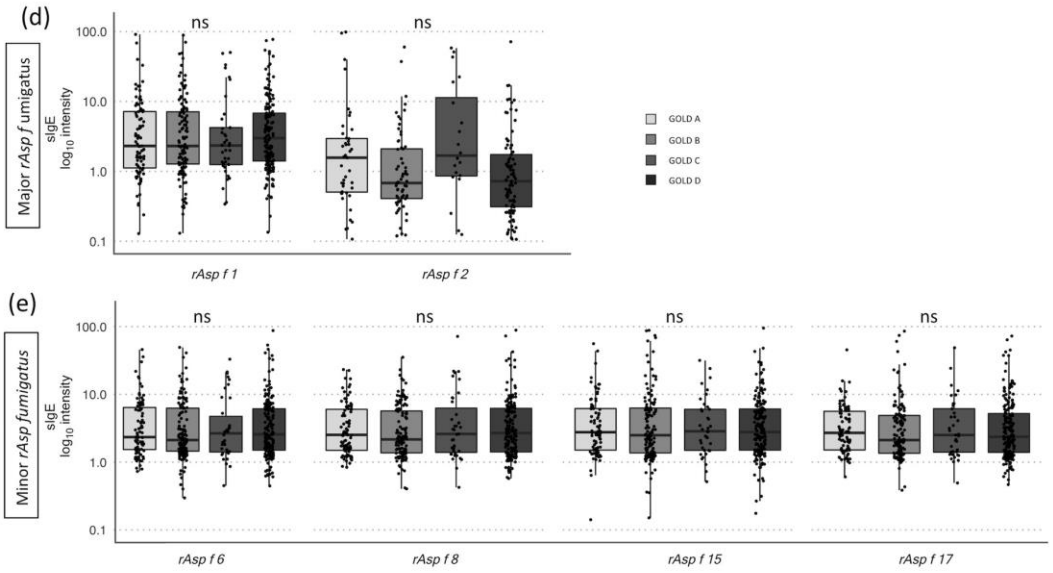

(d)
Major rAsp f umigatus
sIgE
log10 intensity
100.0
10.0
1.0
0.1
ns
ns
rAsp f 1
rAsp f 2
GOLD A
GOLD B
GOLD C
GOLD D
(e)
Minor rAsp fumigatus
sIgE
log10 intensity
100.0
10.0
1.0
0.1
ns
ns
ns
ns
rAsp f 6
rAsp f 8
rAsp f 15
rAsp f 17

Figure E5: Scattered box plots illustrating absolute eosinophil counts (expressed as x 10^9/L) between the detected patient clusters. Colour denotes cluster membership: HS fungal predominant (turquoise), LS (blue), MS *Blomia* predominant (violet-red). HS: high-sensitization; LS: low-sensitization; MS: moderate-sensitization; ns: non-significant. Dotted lines indicates standard eosinophil count cut-off (0.45 x 10^9/L) (grey line) and COPD GOLD cut-off (0.30 x 10^9/L) (red line) to indicate elevated total eosinophil count.

Figure E5

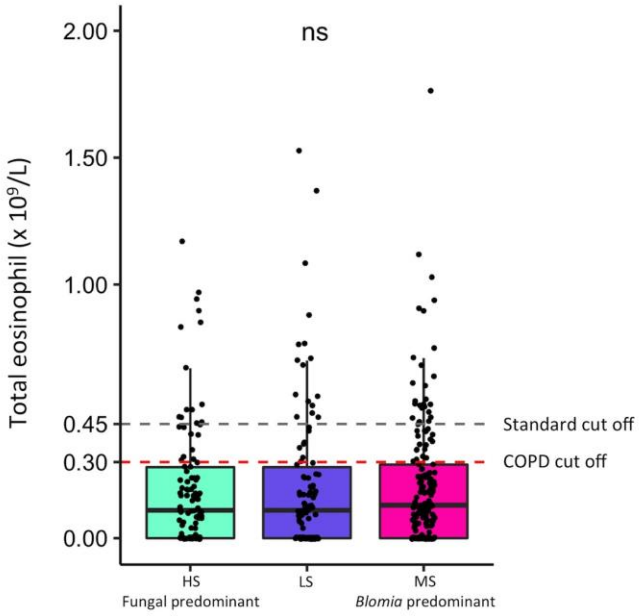


Figure E6: Comparison of (a) symptoms (CAT score), (b) lung function (FEV_1 % predicted), (c) total blood eosinophil count (as 10^9/L) and (d) total serum IgE between non-sensitized and air-fungi sensitized Singaporean COPD patients. CAT: COPD assessment test, FEV_1: forced expiratory volume in the 1st second, IgE: immunoglobulin-E, ns: non-significant.

Figure E6