Sarcoidosis lung microbiota with respect to radiographic types and host genotype

Alexandra Zimmermann, Henrik Knecht, Robert Häsler, Gernot Zissel, Karoline I. Gaede, Sylvia Hofmann, Almut Nebel, Joachim Müller-Quernheim, Stefan Schreiber, Annegret Fischer

Supplementary material

Figure E1. Beta diversity as cluster analysis of sarcoidosis radiographic types, IPFs and healthy controls by 3D-tb-RDA.

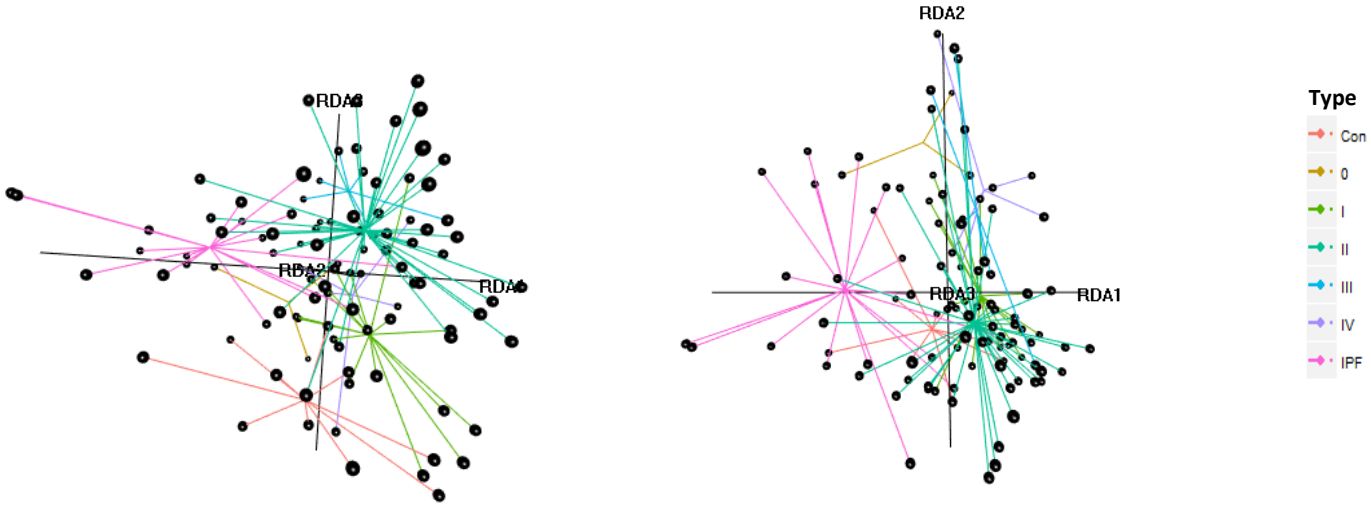

Figure E2. Amount of human DNA versus amount of bacterial DNA in BAL samples (n = 52).

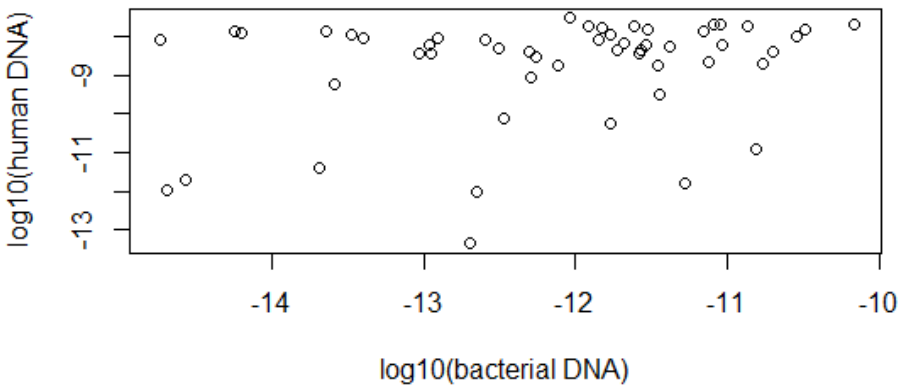

Figure E3. Presence/absence map for the most significant OTUs. The map shows if a specific OTU is present (red) or absent (white) within a sample.

Table E1. Sarcoidosis risk SNPs used for Genotyping

dbSNP ID	Gene	Position	Allele A1/A2	Genotype counts (A1A1/A1A2/A2A2)	Reference
rs12069782	*IL23R*	chr1:67,599,851-67,600,351	C/T	3/19/37	[1]
rs223498	*MANBA*	chr4:103,651,712-103,652,212	A/C	16/23/17	[1]
rs4921492	*IL12B*	chr5:158,832,027-158,832,527	A/C	13/21/26	[1]
rs4143332	*HLA-B*	chr6:31,348,115-31,348,615	C/T	43/14/2	[1]
rs2076530	*BTNL2*	chr6:32,363,566-32,364,066	A/G	23/28/6	[2]
rs5007259	*BTNL2*	chr6:32,378,851-32,379,351	C/T	11/35/12	[1]
rs9277542	*HLA-DPB1*	chr6:33,054,997-33,055,497	C/T	8/22/26	[1]
rs10484410	*ZNF451*	chr6:56,996,438-56,996,938	C/T	49/12/1	[3]
rs1040461	*RAB23*	chr6:57,055,104-57,055,604	T/C	1/7/49	[3]
rs1049550	*ANXA11*	chr10:81,926,452-81,926,952	C/T	4/23/26	[4]
rs479777	*U11*	chr11:64,107,227-64,107,727	C/T	8/14/28	[1]
rs653178	*ATXN2*	chr12:112,007,506-112,008,006	A/G	16/22/15	[1]
rs1050045	*OS9*	chr12:58,115,021-58,115,521	C/T	17/26/13	[5]

Table E2: p-values for alpha-diversity and beta-diversity analyses and bacterial burden stratified by genotype. Alpha-diversity and bacterial burden was tested by pairwise Wilcoxon test. Beta-diversity was tested by ADONIS.

hCV number	Gene	Alpha-Diversity p-value			Beta-Diversity p-value	Bacterial burden p-value		
		A1A1 vs. A1A2	A1A1 vs. A2A2	A2A2 vs. A1A2		A1A1 vs. A1A2	A1A1 vs. A2A2	A2A2 vs. A1A2
hCV8301529	RAB23	1.00	0.25	1.00	0.512	0.89	1.00	0.49
hCV7881261	ANXA11	1.00	1.00	1.00	0.654	1.00	1.00	1.00
hCV2488471	BTNL2	1.00	0.53	0.91	0.862	0.03	<0.01	<0.01
hCV30139071	ZNF451	1.00	1.00	1.00	0.314	1.00	1.00	1.00
hCV2983471	U11	1.00	1.00	1.00	0.426	0.27	0.49	0.27
hCV3188384	OS9	0.06	1.00	0.63	0.654	1.00	1.00	1.00
hCV29999450	MyD88	0.86	1.00	0.34	0.081	1.00	1.00	1.00
hCV25629654	MyD88	1.00	1.00	0.40	0.531	0.88	1.00	1.00
hCV30325393	BTNL2	1.00	1.00	1.00	0.309	1.00	1.00	0.98
rs12069782	IL23R	1.00	1.00	0.82	0.328	1.00	1.00	1.00
hCV27979481	IL12B	1.00	1.00	1.00	0.246	0.24	1.00	1.00
hCV2978544	ATXN2	0.41	0.70	0.68	0.757	0.54	0.16	0.12
hCV26458377	MANBA	0.08	0.25	1.00	0.375	1.00	1.00	1.00
rs4143332	HLA-B	0.06	1.00	1.00	0.135	1.00	1.00	1.00
hCV30508007	HLA-DPB1	0.81	1.0	0.46	0.699	0.12	0.18	0.06

Supplementary methods

Microbial DNA and RNA extraction and preparation

For microbial DNA extraction 1.5 ml of the BAL were centrifuged at full speed for 30 min. The Pellet was affiliated by adding 80 µl of buffer RL of the MolYsis Complete 5 Kit (Molzym GmbH & Co. KG, Bremen, Germany) and further processed according to the kit manual. The MolYsis Complete 5 Kit decreases human bias in bacterial DNA by lysing human cells under chaotropic conditions. The V1-V2 region of the 16S rRNA in the extracted DNA was amplified with primer 27F (5'-GAGTTTGATCCTGGCTCAG-3') combinded with 454 Life Sciences (Roche, Penzberg, Germany) adapter B and 338R (5'-TGCTGCCTCCCGTAGGAGT-3') with 454 Life Sciences adapter A. The reverse primer contained a multiplex barcode identifier sequence (10 bp) which allowed identification of individual samples. The PCR reaction contained 1x Phusion HF Reaction Buffer (1, 5 mM MgCl2), 0.2 µM of the Pyro_27F primer, 200µM dNTPs, 2 U/µl Phusion Hot Start II DNA Polymerase, 0, 2 µM MIDX_338R and 4 µl DNA template. Cycling conditions were 30 Cycles of 10 s denaturation at 98°C, primer annealing for 30s at 57°C and an elongation step for 30s at 72°C, started with an initial denaturation of 3 min at 98°C and terminated with an elongation step of 10 min at 72°C. The samples were prepared for pyrosequencing following the preparation procedure after Stratil et al [6].

Methods for taxonomical analysis

Uncorrected pair-wise distances between the aligned sequences were calculated using the MOTHUR dist.seqs command. To assign the sequences to OTUs the cluster.seqs command with average neighbor method was used [7]. The Shannon-Index was used to quantify the microbial diversity within one sample (alpha diversity [8]) Differences in alpha-diversity and bacterial burden were tested with Wilcoxon rank sum test. To assess the differences in the species distribution between sarcoidosis samples, IPFs and healthy controls, the Kruskall-Wallis test was conducted. Redundancy Analysis (RDA) was carried out on Hellinger transformed data (tb-RDA) [9] with subsequently ADONIS to reveal differences in the OTU compositions between different sample groups. The results for all analyses are corrected for multiple testing by FDR correction. To find sarcoidosis specific species, an indicator species analysis was executed with the R indicspecies package. Thresholds for a valuable indicator species were indicator value > 0.8 and p-value < 0.05.

Supplementary References

1. Fischer A, Schmid B, Ellinghaus D, Nothnagel M, Gaede KI, Schürmann M, Lipinski S, Rosenstiel P, Zissel G, Höhne K, Petrek M, Kolek V, Pabst S, Grohé C, Grunewald J, Ronninger M, Eklund A, Padyukov L, Gieger C, Wichmann H-E, Nebel A, Franke A, Müller-Quernheim J, Hofmann S, Schreiber S. A novel sarcoidosis risk locus for Europeans on chromosome 11q13.1. *Am. J. Respir. Crit. Care Med.* 2012; 186: 877–885.

2. Valentonyte R, Hampe J, Huse K, Rosenstiel P, Albrecht M, Stenzel A, Nagy M, Gaede KI, Franke A, Haesler R, Koch A, Lengauer T, Seegert D, Reiling N, Ehlers S, Schwinger E, Platzer M, Krawczak M, Müller-Quernheim J, Schürmann M, Schreiber S. Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat. Genet.* 2005; 37: 357–364.

3. Hofmann S, Fischer A, Till A, Müller-Quernheim J, Häsler R, Franke A, Gäde KI, Schaarschmidt H, Rosenstiel P, Nebel A. A genome-wide association study reveals evidence of association with sarcoidosis at 6p12. 1. *Eur. Respir. J.* 2011; 38: 1127–1135.

4. Hofmann S, Franke A, Fischer A, Jacobs G, Nothnagel M, Gaede KI, Schürmann M, Müller-Quernheim J, Krawczak M, Rosenstiel P. Genome-wide association study identifies ANXA11 as a new susceptibility locus for sarcoidosis. *Nat. Genet.* 2008; 40: 1103–1106.

5. Hofmann S, Fischer A, Nothnagel M, Jacobs G, Schmid B, Wittig M, Franke A, Gaede KI, Schürmann M, Petrek M. Genome-wide association analysis reveals 12q13. 3–q14. 1 as new risk locus for sarcoidosis. *Eur. Respir. J.* 2013; 41: 888–900.

6. Stratil SB, Neulinger SC, Knecht H, Friedrichs AK, Wahl M. Temperature-driven shifts in the epibiotic bacterial community composition of the brown macroalga Fucus vesiculosus. *Microbiologyopen* 2013; 2: 338–349.

7. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ. Introducing mothur: open-source, platform-independent, community-

supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 2009; 75: 7537–7541.

8. Shannon CE. A mathematical theory of communication, bell System technical Journal 27: 379-423 and 623–656. *Math. Rev. MathSciNet MR10 133e* 1948; .

9. Borcard D, Gillet F, Legendre P. Numerical ecology with R. Springer Science & Business Media; 2011.