

Matched Analysis of Differences of Microbiomes in Japanese vs. Non-Japanese Patients from an Ulcerative Colitis Study

Davit Sargsyan^{1,2}, Javier Cabrera², Kathy Mutambanengwe³,
Dhammika Amaratunga², Ling-Yang Hao¹, Gerold Bongers¹,
Nobutaka Suzuki¹, Lani San Mateo¹, Luc Bijens¹

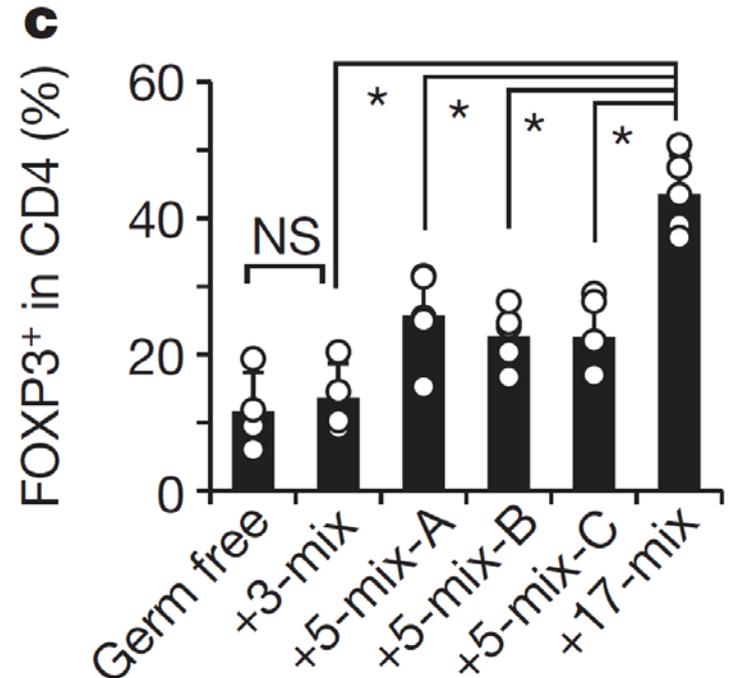
¹Janssen PRD; ²The Cardiovascular Institute of NJ, Rutgers
University; ³Open Analytics

October 2018



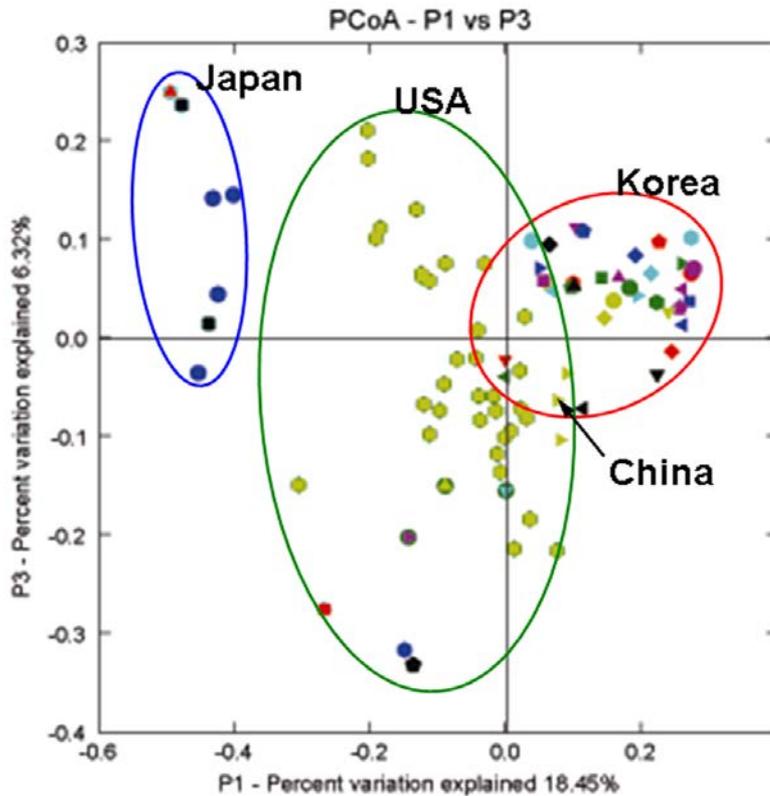
Why Study the Microbiome?

- Gut bacteria interacts with gut cells including immune cells
- Composition of gut bacteria has been shown to modulate inflammation by acting on regulatory T-cells (Atarashi 2013)
- Dysbiosis has been linked to IBD (Chron's disease and ulcerative colitis) (Tamboli 2004)
- Brain-gut-microbiota interaction can affect host's mood and behavior in healthy women (Tillisch 2017)
- Possibly linked to obesity and diabetes



Percentages of FOXP31 cells within the CD41 cell population in IQ1 exGF mice colonized with the indicated mix (Atarashi 2013)

Population Differences in uBiome Composition



Comparison of Korean gut microbial communities to that of non-Korean people (Nam 2011)

- Microbiota can be affected by diet, e.g. higher abundance of *Bifidobacterium* in fecal sample after bread diet compared to rice diet (Mano 2018)
- Culinary preferences and environmental factors may account for uBiome differences between countries and ethnic groups.
- Japanese cuisine and extensive use of seafood is thought to result in distinct microbiota even when compared to neighboring countries (Nam 2011)

Our Study Motivation and Background

- In 2017 Dr. Nobu Suzuki hypothesized finding differences in microbiomes of Japanese and non-Japanese patients
- We considered multiple data sources including publically available data in MetaHIT (EU) and Human Microbiome Project (US) however number of Japanese patients in the available datasets was low (13 and 12 respectively(?))
- We came across Janssen internal data from two studies (C0524T-17 and C0524T-18), from PURSUIT clinical trial

Program of **U**lcerative Colitis **R**esearch **S**tudies **U**tilizing an **I**nvestigational **T**reatment (PURSUIT)

- Ran between August 2007 and October 2010 by Janssen Research and Development (JRD)
- Two Phase 2/3 multicenter, randomized, placebo-controlled, double-blinded studies (C0524T-17 and C0524T-18)
- To evaluate the safety and efficacy of golimumab (Simponi®) in subjects with moderately to severely active ulcerative colitis (UC) as defined by a Mayo score of 6 to 12
- 1,065 UC patients, 18 years of age or older, were randomized to either placebo or high, medium or low dose of treatment: placebo, 400mg/200mg and 200mg/100mg respectively, 331 in each group, with additional 72 subjects randomized to a low dose of 100mg/50 mg, administered subcutaneously (SC)

Data Collection and Processing

- Fecal samples were collected at baseline (Week 0) and 996 of the samples analyzed for bacterial content using 16S ribosomal RNA (rRNA) sequencing
- Out of 996 patients, 71 were Japanese (from Japan)
- Majority of non-Japanese patients were Caucasian (814)
- The largest Asian population after the Japanese were Indians with 44 patients
- The US had the largest number of patients, with 172 Caucasians, 19 Blacks, 2 Asians and 11 identified as other race.

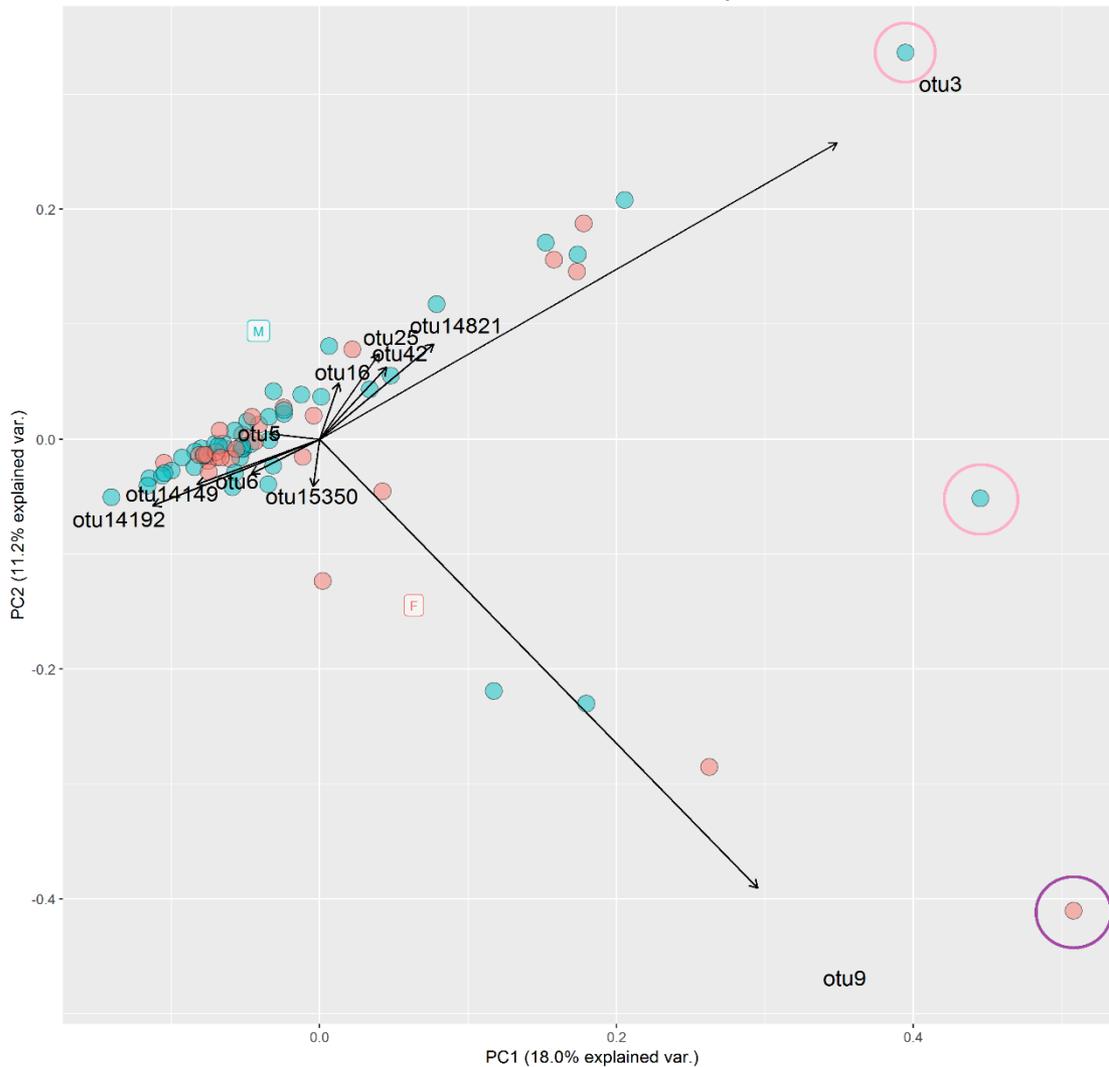
Country	Race				Total
	Asian	Black	Caucasian	Other	
AUSTRALIA	3	0	52	3	58
AUSTRIA	0	1	16	0	17
BELGIUM	0	0	5	0	5
BULGARIA	0	0	53	0	53
CANADA	2	0	36	5	43
CZECH REPUBLIC	0	0	10	0	10
DENMARK	1	0	21	0	22
FRANCE	0	0	16	1	17
GERMANY	0	0	63	1	64
HUNGARY	0	0	81	0	81
INDIA	44	0	0	0	44
ISRAEL	0	1	29	0	30
JAPAN	71	0	0	0	71
NETHERLANDS	0	0	11	0	11
NEW ZEALAND	0	0	15	1	16
POLAND	0	1	107	0	108
ROMANIA	0	0	14	0	14
RUSSIAN FEDERATION	0	0	17	4	21
SERBIA	0	0	19	0	19
SLOVAKIA (Slovak Republic)	0	0	39	0	39
SOUTH AFRICA	0	0	4	5	9
SWEDEN	0	0	2	0	2
UKRAINE	0	0	32	6	38
UNITED STATES	2	19	172	11	204
TOTAL	123	22	815	37	996

Patients' Baseline Characteristics

	Japanese	Non-Japanese
Number of Subjects	71	925
Female (%)	39.4	44.2
Mean Age +/- SD	38.9+/-13.4	40.2+/-13.3
BMI +/- SD	21.6+/-3.3	25.3+/-5.3
Duration of UC (%)		
<=5 years	52.1	56.9
between 5 and 15 years	35.2	35.4
>15 years	12.7	7.8
Extensive/Limited (%)	52/48	59/41
Smoking (%)		
Nonsmokers	60.6	65.2
Prior Smokers	35.2	27.6
Current Smokers	4.2	7.2

- On average, compared to non-Japanese, the Japanese patients were:
- Younger (40.2+/-13.3 vs. 38.9+/-13.4 y.o.)
- Had lower body mass index (BMI of 25.3+/-5.3 vs. 21.6+/-3.20)
- Had a smaller proportion of female (44.2% vs. 39.4%), a smaller proportion of current smokers (7.2% vs. 4.2 %) but a higher proportion of prior smokers (27.5% vs. 35.2%), and a higher proportion of patients with longer duration of UC (7.8% vs. 12.7% of patients with the duration over 15 years).

Relative Abundance Biplot

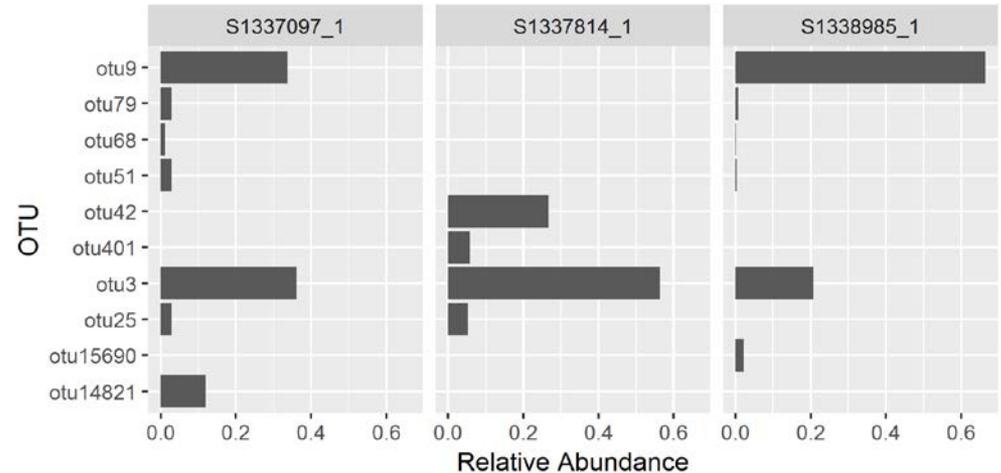


Relative Abundance: PCA of Japanese Patients

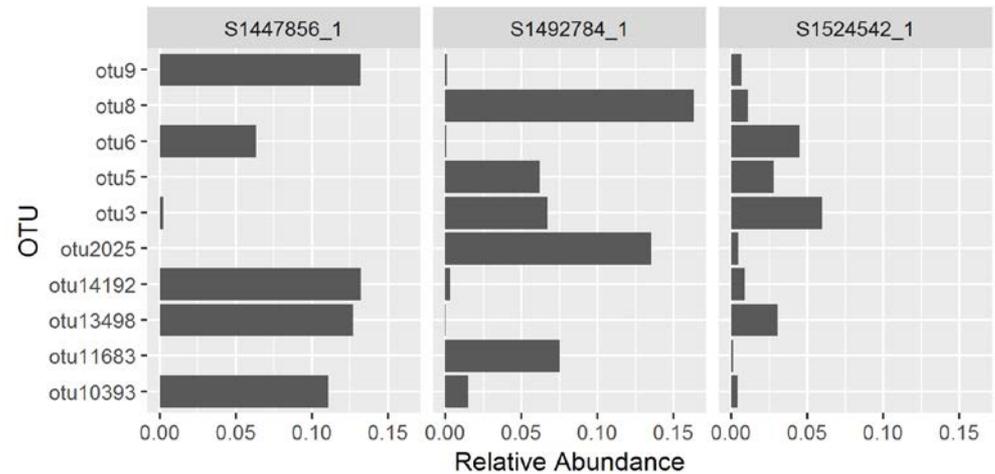
- Some of the patients had microbiome compositions radically different from the majority
- The 3 patients that were the farthest from the majority (right side of the plot, circled in pink, purple and green) had majority of their microbial genetic material mapped to two OTUs – **otu3** and **otu9**

Top 10 OTUs in Most and Least Diverse Samples

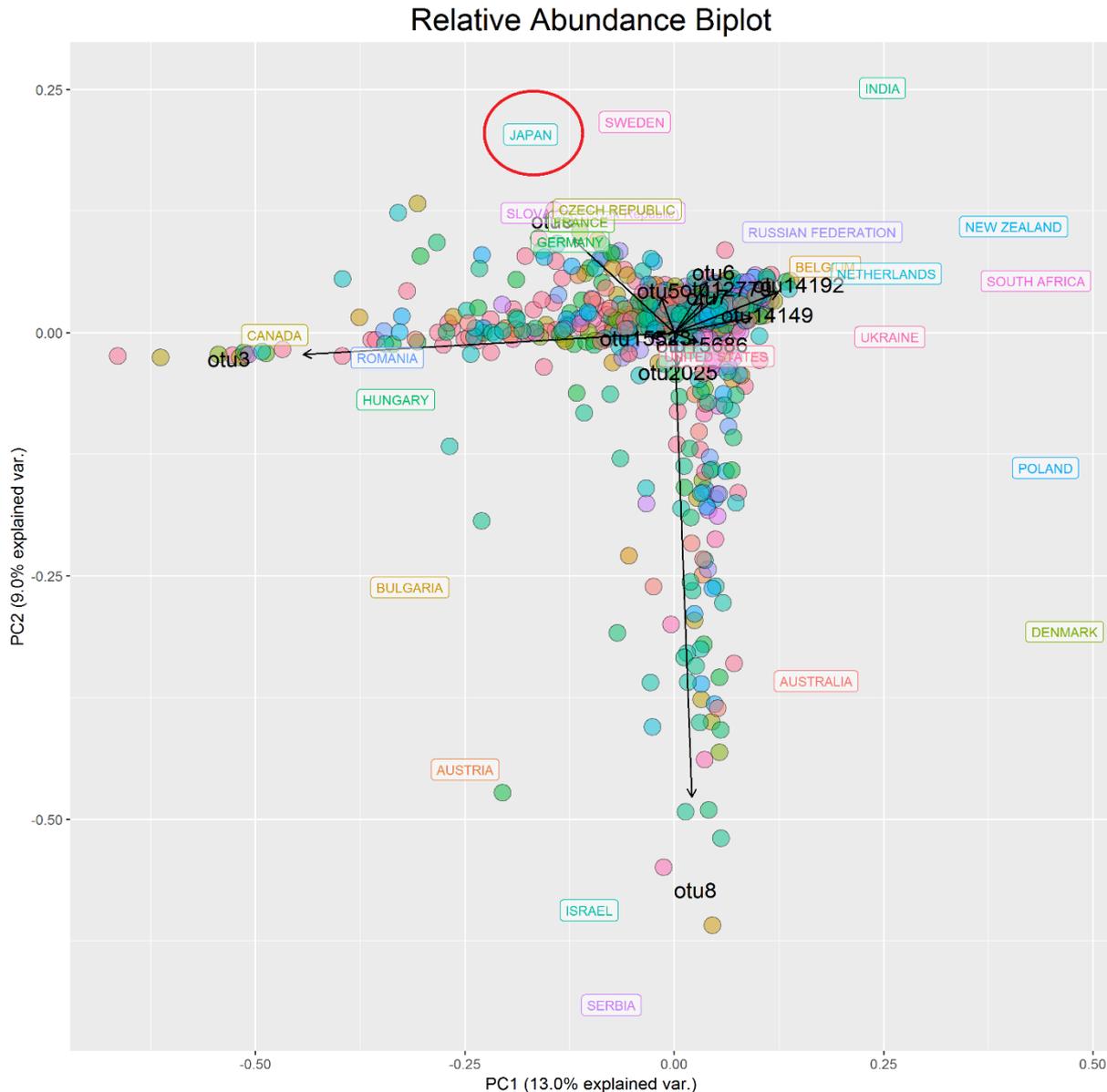
➤ In the extreme samples, just a handful of OTUs accounts for most of the community. E.g. OTU 3 & 9 are up to 60% of these samples (top figure)



➤ The samples that were close to the center of biplot have more even distribution with most abundant OTUs just over 15% (bottom figure)



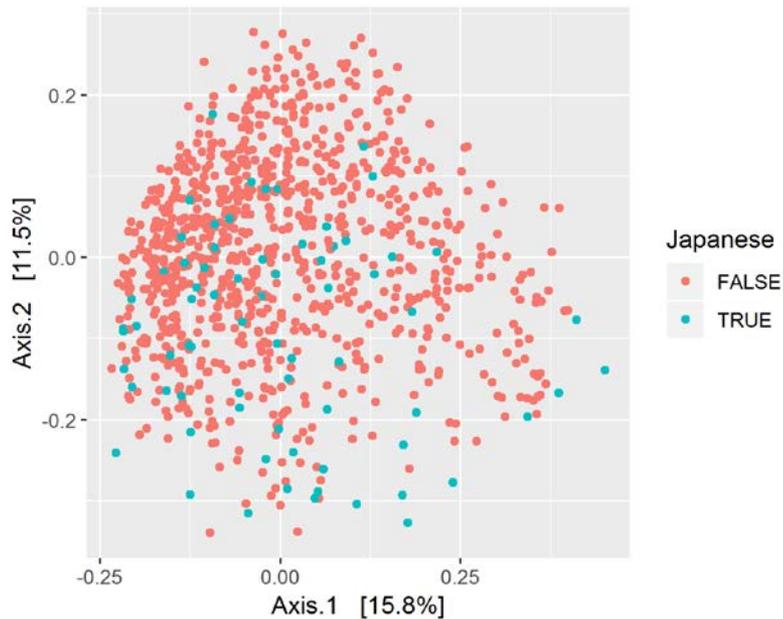
Distribution by Country



- Although no clear separation between the countries can be seen, OTUs 3, 8 and 9 are still influential
- Japanese patients are most dissimilar with the patients from Poland and Denmark, Israel and Serbia
- Japanese patients are most similar to patients from Germany, France, Slovakia, Czech Republic and Sweden.

Jensen-Shannon Divergence (JSD) and Multidimensional Scaling (MDS)

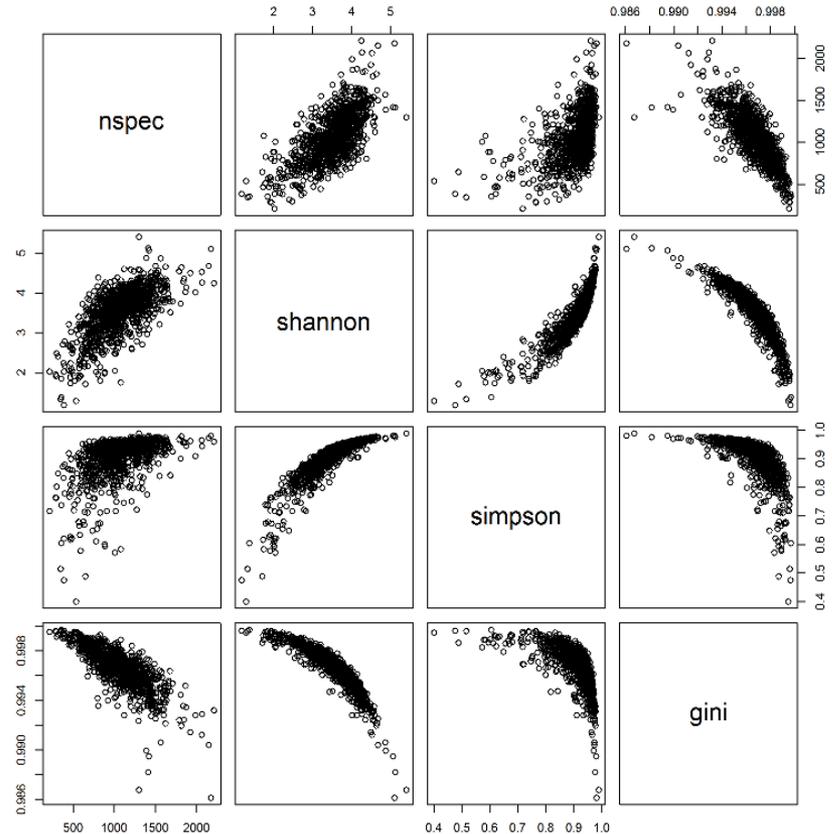
- JSD measures similarities between distributions
- For two vectors of probabilities \mathbf{p} and \mathbf{q} (each adding up to 1), define $\mathbf{m} = 0.5 * (\mathbf{p} + \mathbf{q})$
- $JSD = 0.5 * (\text{sum}(\mathbf{p} * \log(\mathbf{p}/\mathbf{m})) + \text{sum}(\mathbf{q} * \log(\mathbf{q}/\mathbf{m})))$



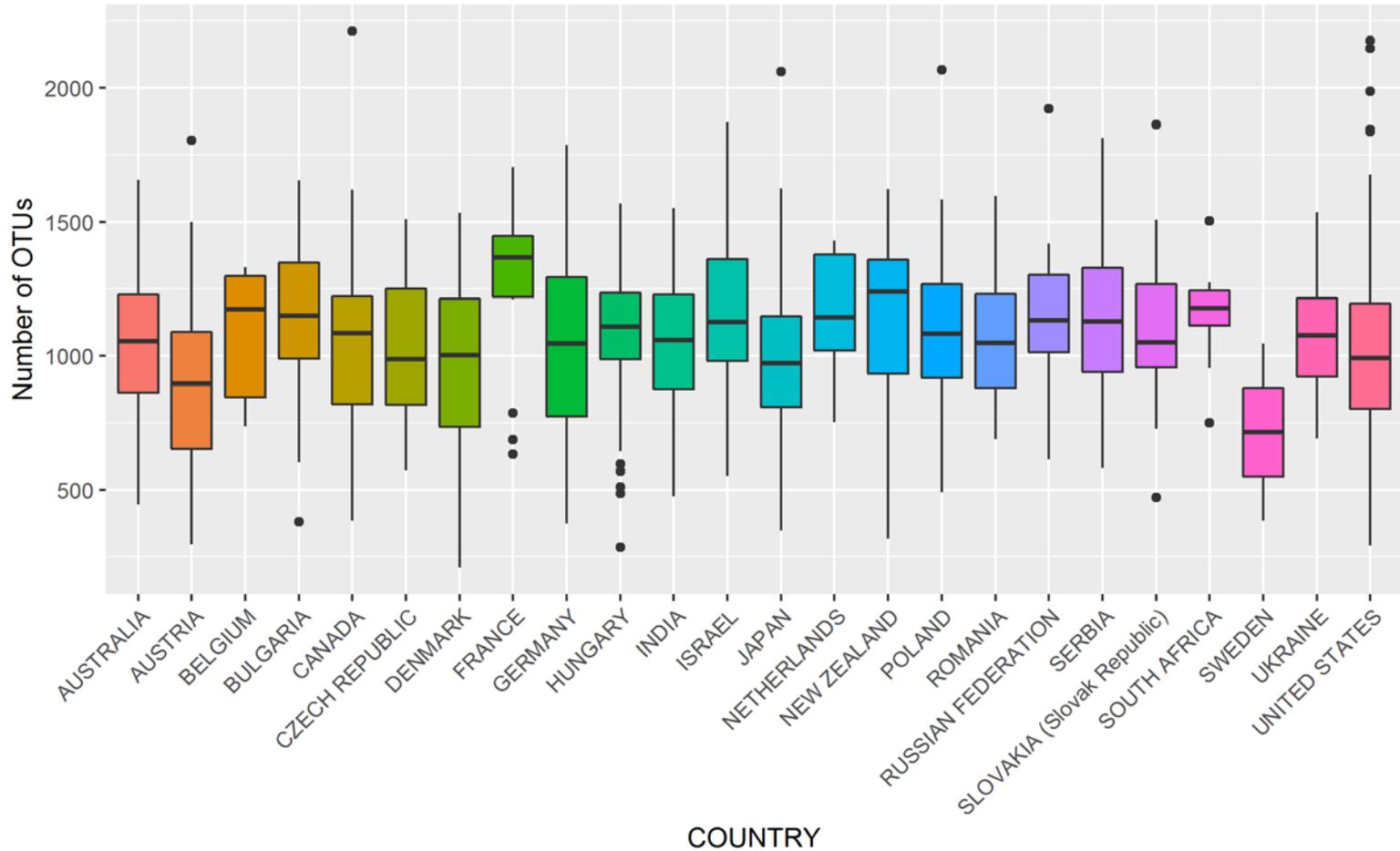
No differences between Japanese and non-Japanese patients using MDS were observed

Richness/Diversity

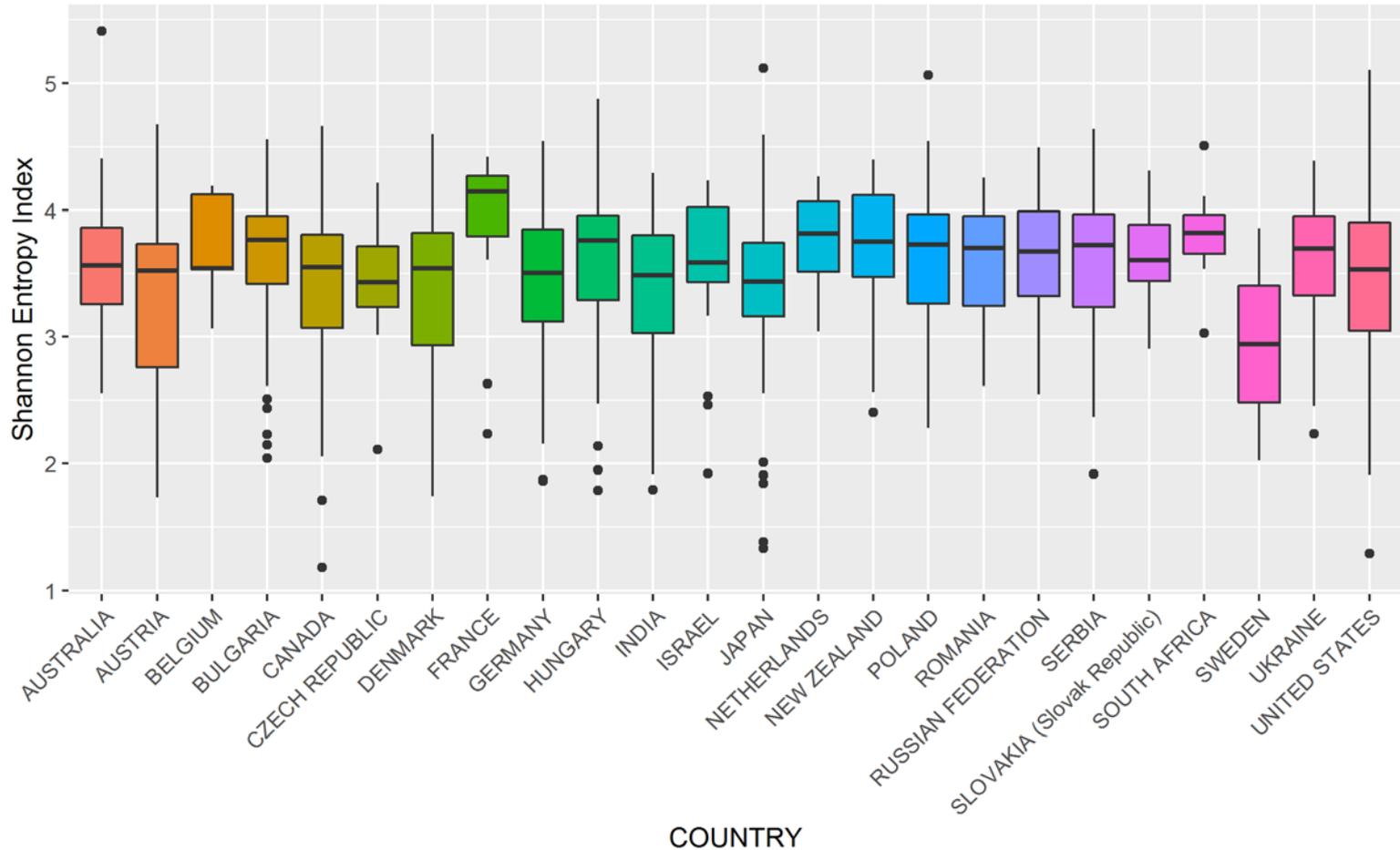
- The simplest metric we calculated was the number of OTUs detected in each sample
- Shannon entropy index $H' = -\sum_{i=1}^R p_i \ln(p_i)$, where p_i is the proportion of the i -th species in the sample, and R is the total number of species present in the sample
- All 4 metrics correlate



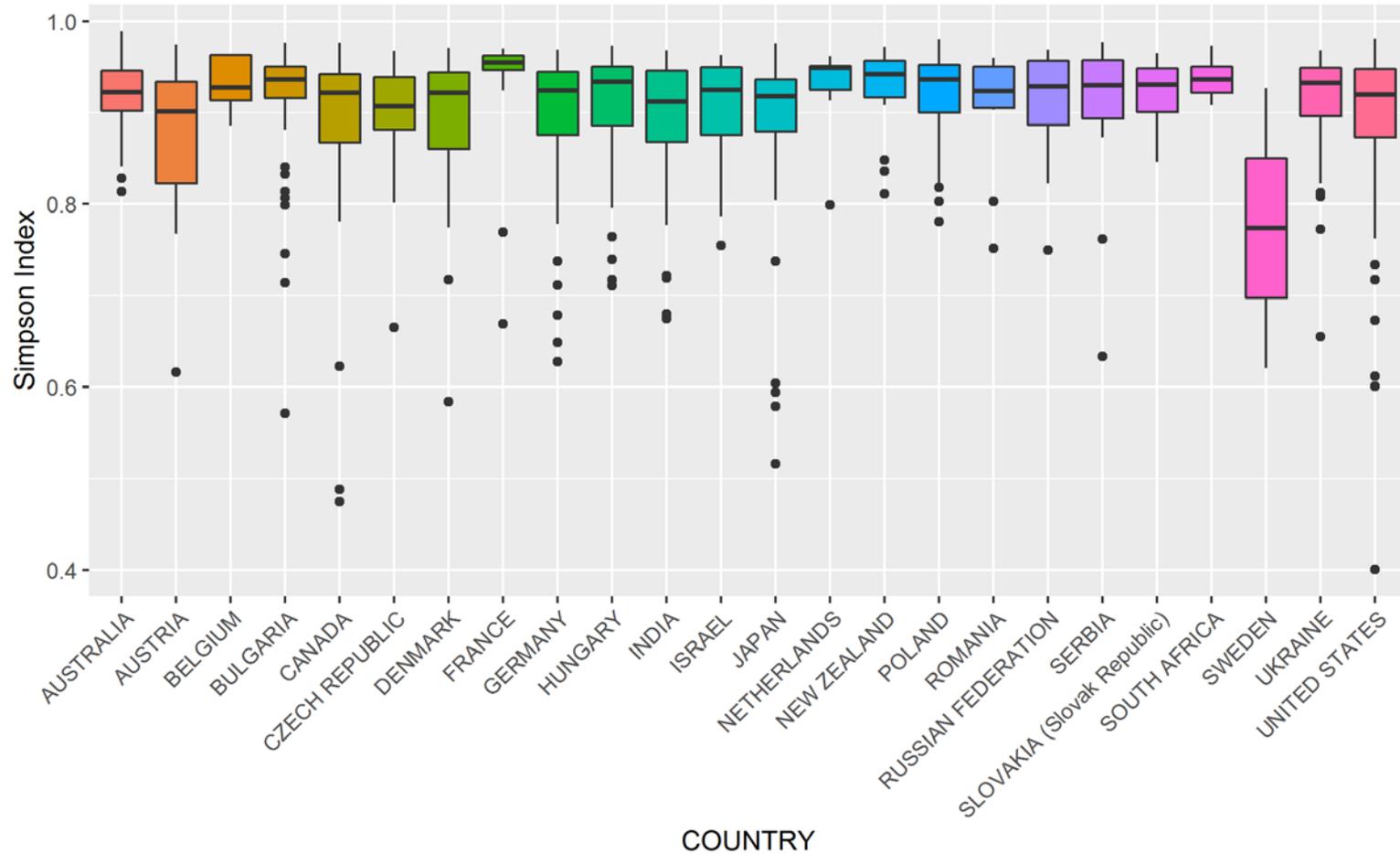
Richness by Country



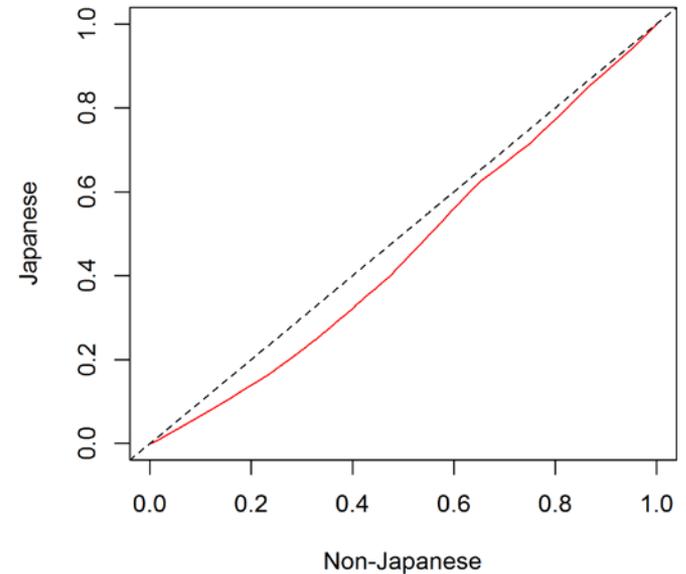
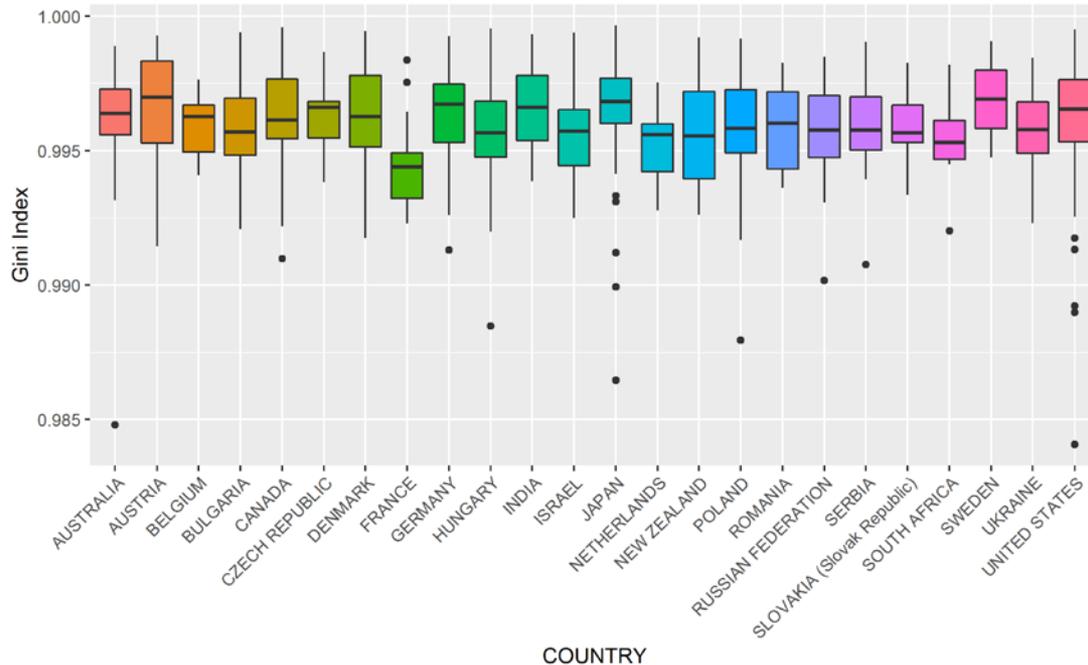
Shannon Entropy by Country



Simpson Index by Country



Gini Coefficient by Country



- On average there was more inequality between the OTUs in Japanese patients compared to non-Japanese (larger Gini coefficient)
- The cumulative sum of relative abundances rose slower in Japanese patients (more inequality)

Modeling Results

- At OTU level, bacterial diversity was lower in Japanese patients compared to non-Japanese.
- In the unmatched dataset the average number of detected species were 996.8 ± 35.2 and $1,073.8 \pm 9.7$, corresponding to Shannon indices of 3.38 ± 0.08 and 3.53 ± 0.02 for Japanese and non-Japanese patients respectively.
- In the univariable model with Shannon index as a response and ethnicity (Japanese/non-Japanese) as the main effect, the difference was statistically significant (Shannon index difference of 0.15 ± 0.07 , p-value = 0.034), and it remained significant after adjusting for patient characteristics (p-value = 0.045).

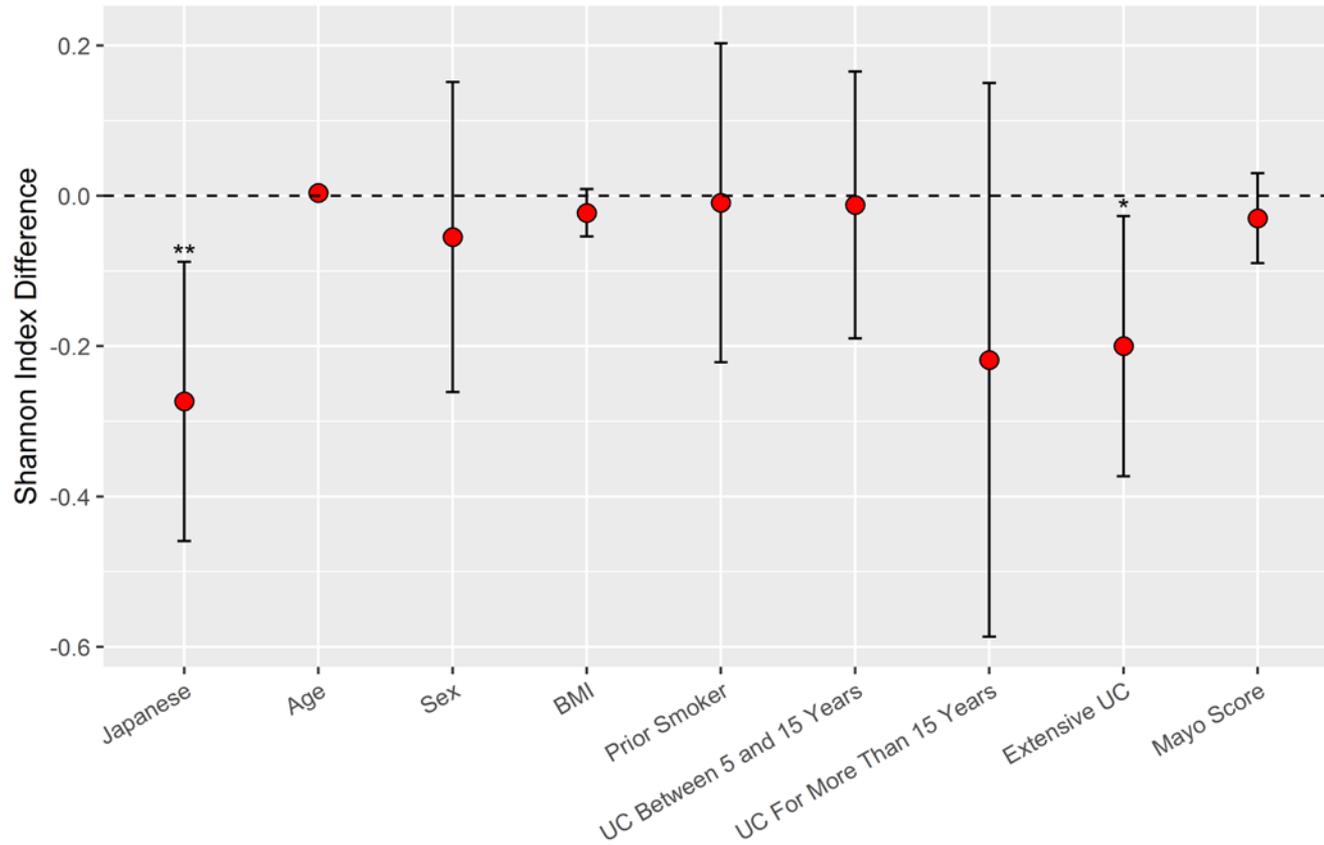
*All estimates are reported as *mean +/- standard error of the mean (SEM)*

**Estimates from a univariable linear model

Matched Analysis

- We matched each Japanese patient with up to three non-Japanese patient using exact matching for sex, smoking status, duration and extent of the disease, and approximate matching (within a delta) for age, BMI and Mayo score
- By adjusting deltas in the approximately matched covariates, we leveraged tighter match against number of matched patients. For example, constraining age and BMI differences within 5 years, and Mayo score within 2 points, 58 out of 71 Japanese patients were matched to 129 out of 925 non-Japanese. This produced a Shannon index difference of 0.22 ± 0.09 . The difference was significant in both, univariable and multivariable models (p-values of 0.016 and 0.004 respectively).
- Alternatively, estimates from a stratified linear model were 0.26 ± 0.09 . There was no statistically significant difference between the two models.
- The difference was significant in both, univariable and multivariable models (p-values of 0.016 and 0.004 respectively).

Mixed-Effects Linear Model with Covariates Matched Dataset



References

1. *Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota*, Atarashi et al, Nature Letters, August 2013
2. *Brain structure and response to emotional stimuli as related to gut microbial profiles in healthy women*, Tillisch et al, Psychosom Med, October 2017
3. *Dysbiosis in inflammatory bowel disease*, Tamboli et al, Gut Journal, 2004
4. *The Effect of White Rice and White Bread as Staple Foods on Gut Microbiota and Host Metabolism*, Mano et al, Nutrients, 2018
5. *Comparative Analysis of Korean Human Gut Microbiota by Barcoded Pyrosequencing*, Nam et al, PLoS ONE, July 2011