# Supplementary material

**Table S1 The development of ovaries in each group (n=5)**

| | CON | EGCG | HFD | HFDEGCG |
|---|---|---|---|---|
| Primordial follicle (n) | 13.20 ± 1.98 | 8.40 ± 2.40 | 10.25 ± 1.03 | 13.80 ± 2.42 |
| Primary follicle (n) | 8.40 ± 0.51 | 7.40 ± 1.44 | 8.25 ± 1.70 | 8.60 ± 1.66 |
| Secondary follicle (n) | 2.40 ± 0.87 | 1.00 ± 0.71 | 1.50 ± 0.65 | 2.20 ± 0.49 |
| Corpus luteum (n) | 5.40 ± 0.51 | 3.00 ± 0.63 | 4.67 ± 1.20 | 3.60 ± 0.40 |

**Calculation formula of GMHI** :

$$h_{i,M_H,M_N} = \log_{10}\left(\frac{\frac{R_{M_H}}{|M_H|}\sum_{j\in I_{M_H}}\left|n_j ln\left(n_j\right)\right|}{\frac{R_{M_N}}{|M_N|}\sum_{j\in I_{M_N}}\left|n_j ln\left(n_j\right)\right|}\right)$$

Analysis and calculation steps of GMHI [7]:

1) First determine the prevalence of $p_{H,m}$ and $p_{N,m}$ , or microbial species m, in H and N (prevalence is defined as the proportion of samples in a given group, where m is considered "present", i.e. relative abundance ≥$1.0\times10^{-5}$).

2) In order to compare the two prevalence rates of H and N, the following two criteria were used: change in prevalence multiple ($f_m^{H,N}$) and difference in prevalence, defined as $\frac{p_{N,m}}{p_{H,m}}$ ,and $p_{H,m}-p_{N,m}$,respectively.

3) If both criteria meet the minimum threshold for the (predetermined) prevalence multiple change θf and the prevalence difference θd (how do we determine the optimal threshold pair), then a significant effect size is considered to exist between the two prevalence rates. For all detectable microbial species that meet both $f_m^{H,N}$ ≥θf (multiple range 1.2-2.0, minimum θf if effect size values are consistent) and $d_m^{H,N}$ ≥θd, we assign these values to H (more than N). Species that are observed more frequently are called "healthy epidemic $M_H$" species. Similarly, we identified "health-scarce" species $M_N$.

4) In this respect, the species ultimately selected to make up MH and MN both depend on θf and θd. An important advantage of our prevalence based strategy for identifying microbial associations is that it does not calculate or compare the average of measurements from different sources, which makes it difficult to demonstrate that biological and technical heterogeneity may vary widely.

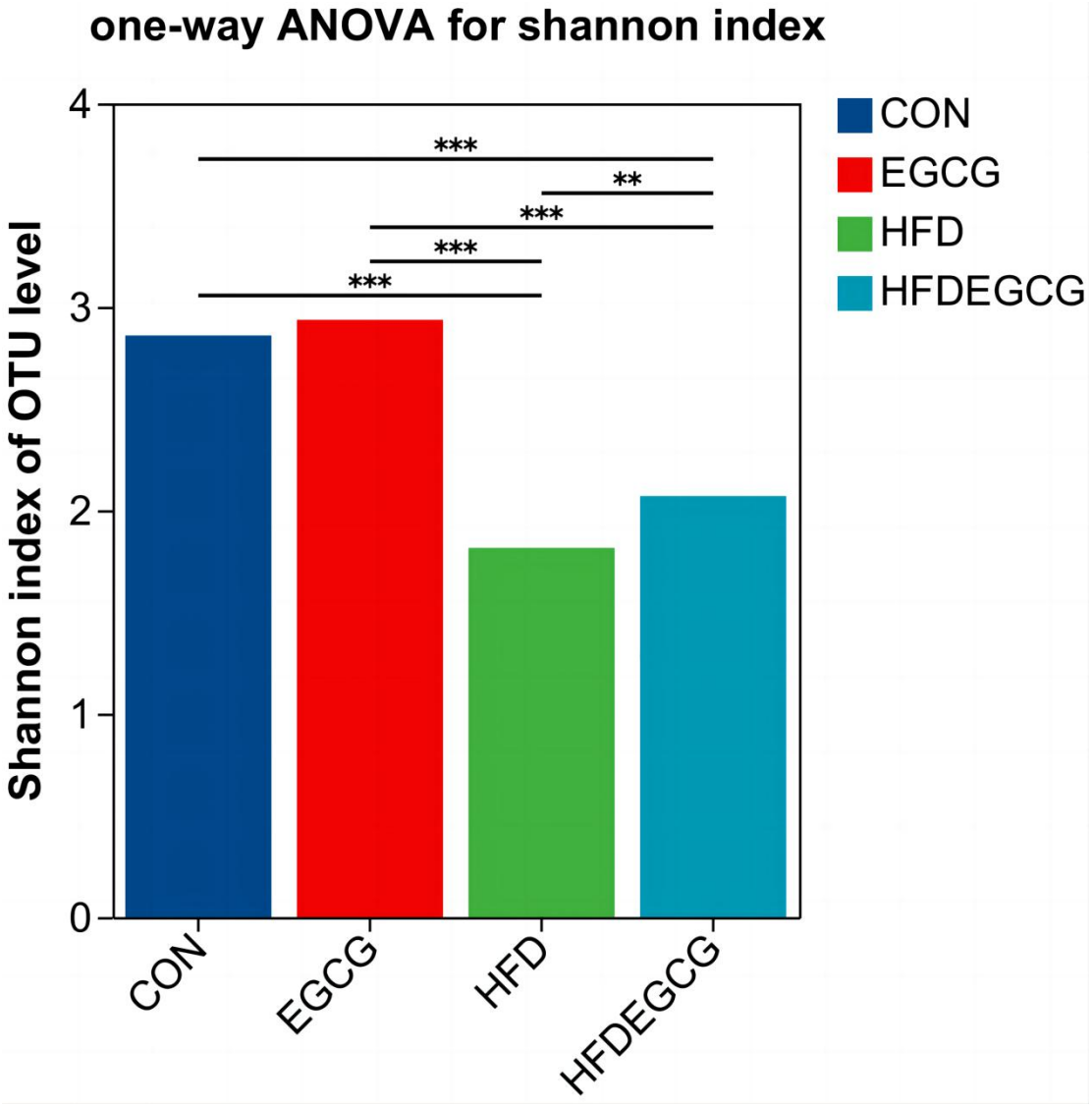


**Figure S1 The Shannon index of OTU level for four groups.**

Data are presented as mean ± SEM (n=5/group). **P* < 0.05, ***P* < 0.01, and ****P* < 0.001.