A theoretic approach to the breast tissue microbiomes under tumor influences with near-neutral, neutral and niche-neutral hybrid models

Lianwei Li\textsuperscript{1,3}  Zhanshan (Sam) Ma\textsuperscript{1,2,3*}

\textsuperscript{1}Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

\textsuperscript{2}Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

\textsuperscript{3}Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China

*For all correspondence to: ma@vandals.uidaho.edu

Abstract

Background: The structure and dynamics of breast tissue microbiomes can have far-reaching influences on women’s health, particularly on breast tumor development. However, there is currently little understanding on the ecological processes that shape the structure and dynamics of breast tissue microbiomes.

Methods: Here we fill the gap by applying three metacommunity models for investigating the community assembly and diversity maintenance, including Sloan near neutral model, Harris et al. multisite neutral and Tang & Zhou niche-neutral hybrid models to reanalyze the 16s-rRNA sequencing datasets of 23 healthy, 12 benign tumor, and 33 malignant tumor tissue samples. To remedy the practical difficulty in collecting tissue microbiome samples, we adopted a sophisticated random re-sampling (up to 1000 times) scheme in applying the three metacommunity models for analyzing the patterns in the breast tissue microbiomes.

Results: First, we found that, at the community/metacommunity levels, the mechanisms of microbiome assembly and diversity maintenance of breast tissue microbiomes were predominantly driven by stochastic drifts of bacteria demography (division, death and dispersal of bacterial cells), whereas the deterministic selection forces such as tumor progression were insignificant. However, at species level, on average, approximately 10\% and 5\% species were above (positively-selected) and below (negatively-selected) neutral, respectively. Furthermore, malignant tumor may raise the positively selected species up to 17\%. Second, malignant tumor appears to inhibit microbial dispersal as evidenced by lowered migration rates, compared with the migration in normal and benign tumor tissues.

Conclusions: The mechanisms of microbiome assembly and diversity maintenance of breast tissue microbiomes were predominantly driven by stochastic drifts, and malignant tumor may inhibit microbial dispersal. These theoretic findings can be inspirational for further investigating the relationships between tissue microbiomes and breast tumor progression/development.
Keywords: Breast tumor; Tissue microbiomes; Species-level neutrality analysis; Multi-site neutral model (MSN); Niche-neutral hybrid model (NNH);

Introduction

With the rapid development of 16S amplicon sequencing, the microbiome of breast milk has become an area of interest for research on the health of infants and mothers. The microbiome in human milk may be the seed source of gut microbiota in infants, with its composition and diversity found to be correlated with body-mass index (BMI), parity, mode of delivery, breastfeeding practices, and infant oral cavities. For example, breastfeeding mode is a key determining factor of milk microbiota composition (Moossavi et al., 2019). Furthermore, microbiomes can influence susceptibility to cancers and the response to therapeutics (Helmink et al., 2019). Many researchers have focused on the relationship between microbial profiles and disease development. Human milk and breast tissue contain microbial communities that are thought to be sterile (Hunt et al., 2011; Urbaniak et al., 2014, 2016; Hieken et al., 2016). Evidence indicates that distinct microbial communities exist among healthy, benign, and malignant breast tissue. Hunt et al. (2011) explored the microbial profiles in human milk based on pyrosequencing of the 16S ribosomal RNA gene and found that the milk microbiome is relatively stable over time within an individual. However, this is not always true. Ma et al. (2015) re-analyzed bacterial interactions using the same datasets as Hunt et al. (2011) and found that dysbiosis of the milk microbiome following a shift in the balance between potential opportunistic pathogens and harmless bacteria is likely responsible for some breast diseases. The gut microbiome can also affect breast cancer due to the estrobolome, i.e., bacterial genes capable of metabolizing estrogens, and thus can affect the emergence of estrogen-driven breast cancers (Kwa et al., 2016; Goedert et al., 2015).

As described previously, breastfeeding and estrogen can impact the breast microbiome profile and diversity. It has been reported that diversity in the milk microbial community is decreased in Hodgkin’s lymphoma cohorts compared with healthy controls, and is positively correlated with milk metabolites that benefit children and negatively correlated with harmful metabolites related to mastitis and breast cancer (Ma et al., 2016). Obviously, the interaction between microbes and breast disease can change the composition and diversity of the microbiome. However, the mechanisms of breast microbiome construction in healthy and diseased groups remain unclear.

To explore the mechanisms of microbial community construction and diversity maintenance, niche theory and neutral theory can be applied to illustrate the role of stochastic and deterministic forces. Hubbell (2001) initially applied the unified neutral theory of biodiversity and biogeography (UNTB) to explain community construction mechanisms, with many researchers subsequently extending and challenging traditional neutral and niche theories (Etienne, 2005, 2007; Hubbell 2006; Volkov, 2003, 2007; Sloan et al., 2006, 2007; McGill, 2003; Harris et al., 2017; Tang & Zhou, 2013, Li & Ma 2016; Burns 2016). For example, Sloan et al. (2006, 2007) extended Hubbell’s discrete neutral theory to a continuous version to test large microbiomes. This model can define whether a species is neutral or not beyond the community level and can be applied to identify important microbes in the community construction process. In the current study, we applied Sloan’s neutral community model to define neutral and non-
neutral species in healthy (control), benign, and malignant tumor cohorts and to explain the dynamics of neutral species.

Compared with the original Hubbell (2001) UNTB model, Sloan (2006, 2007) model is actually a near neutral model since it allows for the existence of competitive advantages or disadvantages. Thanks to this advance, all species in a community can be categorized as three types: neutral species, negatively selected (under neutral) and positively selected (above neutral). We take advantage of this feature to detect the potential correlation between tumor development and bacterial species competitiveness. We also used another pair of models, *i.e.*, multi-site neutral (MSN) (Harris et al., 2017) and niche-neutral hybrid models (NNH) (Tang & Zhou, 2013), to evaluate the relative significance of stochastic neutral drifts vs. deterministic niche selection in driving community assembly and shaping the diversity patterns of the breast tissue microbiome. The MSN by Harris *et al.* (2017) is a major computational advance to Hubbell’s classic UNTB because it allows for simultaneously estimation of the migrations rates among large number of sites (local communities), which was a significant computational challenge until recently. Obviously, this simultaneous estimation of migration parameters is closer to reality. While the MSN model is an orthodox implementation of Hubbell’s UNTB model, and Sloan model is a near neutral model, the NNH model is a mixture (hybrid) of neutral and niche mechanisms. Therefore, the three models we choose to apply in this study span the whole spectrum of the so-called niche-neutral continuum, which postulates that different types of metacommunities are likely fall in different locations of the continuum. One end of the spectrum is occupied by the completely neutral assemblages, and another end is by completely niche-differentiated assemblages. Therefore, the objective of this article is to examine the ‘position’ of the breast tissue microbiome on the niche-neutral continuum, particularly when tumor development occurs with the breast tissue. The integrated analysis with the three models allows us to present a relatively complete and reliable ‘picture’ of the process (mechanism) underlying the tissue microbiome distribution and dispersal patterns. In perspective, our study, if successful, is likely to offer important insights for investigating the relationship between breast tumor development and breast tissue microbiomes (Nejman et al 2020; Poore et al 2020).

**Materials and Methods**

**Microbiome in normal, benign, and malignant breast tissue**

Urbaniak *et al.* (2016) collected samples from women following lumpectomies or mastectomies and from healthy individuals. The breast tissue microbiome datasets consisted of three groups: *i.e.*, healthy (23 samples), benign tumor (12 samples), and malignant tumor tissues (33 samples). Samples were taken from normal tissue adjacent to the tumor, rather than from the tumor tissue itself. The 16S rRNA sequencing datasets were analyzed by QIIME to obtain the operational taxonomic unit (OTU) tables.

**Sloan (2006, 2007) near neutral model**

Sloan *et al.* (2006, 2007) derived a neutral model to explain the assembly mechanisms of prokaryotic communities. As a continuous version of Hubbell’s discrete neutral community model, Sloan’s model does not require observed species abundance distributions or patterns and
can test very large prokaryotic communities. The model contains source and local communities, similar to “mainland” and “island” in the theory of island biogeography. We can first assume that the local community is saturated with \( N_T \) individuals. One individual dies or leaves the local community and is replaced by another individual immigrating from a source community with probability \( m \) or offspring of a random individual within local community with probability \( 1-m \). Thus, the probability that the abundance of the \( i \)-th OTU increases by one individual, decreases by one individual, or shows no change can be given by:

\[
\begin{align*}
\Pr(N_i + 1/N_i) &= \left(1 - \frac{N_i}{N_T}\right)[mp + (1 - m)\left(\frac{N_i}{N_T} - 1\right)] \\
\Pr(N_i - 1/N_i) &= \frac{N_i}{N_T} \left[m(1 - p_i) + (1 - m)\left(\frac{N_T - N_i}{N_T} - 1\right)\right] \\
\Pr(N_i/N_i) &= \frac{N_i}{N_T} \left[m(1 - p_i) + (1 - m)\left(\frac{N_T - N_i}{N_T} - 1\right)\right]
\end{align*}
\]

where \( p_i \) is the occurrence frequency of the \( i \)-th OTU in the source community and \( N_i \) is the abundance of \( i \)-th OTU in the local community. Let \( x_i = N_i/N_T \) be the occurrence frequency of the \( i \)-th OTU in the local community. The prediction abundance (\( \phi \)) of community is the beta distribution:

\[
\phi_i = CX_i^{N_Tm - 1}(1 - X_i)^{N_Tm(1 - p_i) - 1}
\]

where, \( C = \frac{\Gamma(N_Tm)}{\Gamma(N_Tm(1 - p_i))\Gamma(N_Tm)} \).

From Sloan’s model, we can judge whether each species is neutral or not. According to Burns et al. (2016), the process for testing Sloan’s neutral model can be summarized as follows:

1. Compute \( p_i \) and \( x_i \), fitting beta distribution and obtaining the estimation of \( m \).
2. Compute the theoretical occurrence frequency of species \( i \) across all local community samples with \( m \) and the beta distribution.
3. Judge whether the observed \( x_i \) of species \( i \) falls within its 95% theoretical interval predicted from the neutral community model, and obtain a list of neutral, below neutral, and above neutral species.

**Multi-site neutral (MSN) model**

Harris et al. (2017) is an implementation of Hubbell (2001) unified neutral theory of biodiversity (UNTB) by approximating the multinominal (MN) species abundance distribution model with a hierarchical Dirichlet process (HDP). The derived algorithm can simultaneously estimate the migration rates among reasonable large number of sites, and therefore, we term the model as multi-site neutral model (MSN) or HDP approximated MSN model (HDP-MSN). With MSN modeling, the neutrality test can be performed at both local community and metacommunity level simultaneously. In other words, there are \( P \)-values for local community neutrality and metacommunity neutrality, respectively. For detailed computational procedures and software program of the MSN model, one may refer to Harris et al. (2017).

**Niche-neutral hybrid model**

Tang & Zhou (2013) proposed a hybrid niche-neutral model for multiple discrete communities
developed by Volkov et al. (2007). Volkov et al. (2007) assumed that interspecies interactions in a steady-state community can be ignored and all species in the community become functionally equivalent. Here, based on the datasets of the healthy and breast tumor microbiomes, we treated each sample as a niche occupied by a local microbial community and fit the neutral model for each local community. We used the \( p \)-value of the Chi-squared test to determine whether the metacommunity fit the NNH model. At the metacommunity level, if \( p > 0.05 \), the metacommunity satisfies the NNH and its assembly is co-driven by both niche and neutral processes, implying that the metacommunity itself does not satisfy the neutral theory, but within each niche, the local community is neutral; if \( p < 0.05 \), the metacommunity does not satisfy the NNH, implying that within each niche, the local community is not neutral either, and the metacommunity assembly is solely influenced by the niche process.

**The overall modeling design and computational implementations**

Summarizing previous sub-sections for the breast tissue microbiome datasets, as well as the three metacommunity models, we still need to design computational implementations to apply the models to the datasets for achieving our objectives—assessing and interpreting the relative important of stochastic neutral drifts and deterministic selection (specifically tumor progression in this study). Table 1 below outlines our study design. An important step in our design was the adoption of random re-sampling (reallocations) of the samples for 1000 times from each of the three tissue categories (normal, benign and malignant) or 100 times from each of the pair-wise two categories (to evaluate the directional changes). The re-sampling of 1000 (100) times was used build metacommunities, and therefore 1000(100) metacommunity models (Sloan, MSN or NNH) were built for the whole datasets. The random re-sampling was used to raise the robustness of the modeling, in particular to compensate for the relative small sample sizes. Obviously, the tissue microbiome samples are much more difficult to obtain than those in non-invasive studies such as stool or oral microbiome sample collections.

**Table 1.** The study design for testing Sloan (2006, 2007) near neutral, Harris et al. (2017) neutral and Tang and Zhou (2013) niche-neutral hybrid models for three types of breast tissue microbiome samples

| Scale              | Sampling Procedures                                                                 | Models                                                                 |
|--------------------|-------------------------------------------------------------------------------------|------------------------------------------------------------------------|
| **Species Level**  | Select a group of community samples as the source community, another group as the destination (local) community, i.e., (i) Normal (source) to Benign (local) (ii) Normal (source) to Malignant (local) (iii) Benign (source) to Malignant (local) | Sloan’s near neutral model (Sloan et al 2006, 2007)                      |
| **Metacommunity Level** | Randomly select one community sample from each of the three groups: 23 normal tissue samples, 12 benign tissue samples, and 33 malignant tissue samples. The sampling was performed with replacement, there were \( 23 \times 12 \times 33 = 9108 \) possible combinations (metacommunities). We randomly select 1000 out of the 9108 without replacement and fit the model to each of the 1000 metacommunities. | MSN (Harris et al 2017), NNH (Tang & Zhou 2013)                          |
|                    | Randomly select one community sample from each of the three pairs of groups, i.e., \{normal, benign\} \{normal, malignant\} and \{benign, malignant\}, to form a metacommunity of two local communities. For each of the following three types of metacommunities, only 100 times of re-sampling were performed. That is, (i) From the combinations of the 23 normal tissue samples 12 benign samples, there were \( 23 \times 12 = 144 \) possibly combinations (metacommunities), 100 metacommunities are randomly selected from the 144 possible combinations. Similarly, 100 metacommunities are selected from the following combinations: (ii) Normal (23) and Malignant (33), \( 23 \times 33 = 759 \) (iii) Benign (12) and Malignant (33), \( 12 \times 33 = 396 \) | MSN (Harris et al 2017), NNH (Tang & Zhou 2013)                          |
Results and Discussion

Identifying neutral, negatively selected (below-neutral) and positively selected (above-neutral) species with Sloan (2006, 2007) near neutral model

Using Sloan’s model, all species in a metacommunity can be classified into three groups: neutral, below neutral (negatively selected) and above neutral (positively selected). As shown in Table 2, Table S5A and Figure 1, when we treated the normal tissue microbiomes as source community and the benign tumor ones as destination community, 84.8% of the species belonged to neutral category, 5.6% were negatively selected species, and 9.6% were positively selected species. Given that the normal tissue microbiomes were treated as source community, when we treated the malignant tumor ones as destination community, 76.5% of the species were driven by stochastic neutral forces, 8% were driven by negatively selected, and 15.5% were driven by positively selected (Table 2, Table S5B & Figure 2). We also tested Sloan’s model by treating the benign tumor tissue microbiomes as the source community and the malignant tumor ones as the destination community, in which 75.2% of the species satisfy neutral theory, 8% belonged to negatively selected species, and 16.8% belonged to positively selected species (Table 2, Table S5C & Figure 3).

From these results and the results of significance tests of difference, we found that the percentage of neutral species in tissue microbiome significantly decreased with the transformation of tumor from benign to malignant; meanwhile, the percentage of non-neutral species, especially positively selected species, significantly increased when malignant transformation occurred (Fisher test: $p$-values < 0.001; Figure 4). Specifically, in the breast tissue microbiome, approximately 10% of the species is positively selected by the progress of benign tumorigenes, but the percentage of the species positively selected by malignant transformation is up to about 17%.

| Source community | Destination community | $N$   | $m$   | $R^2$ | Total | Neutral (%) | Below neutral (%) | Above neutral (%) |
|------------------|-----------------------|-------|-------|-------|-------|-------------|------------------|------------------|
| Normal           | Benign                | 15685.667 | 0.003 | 0.185 | 712   | 84.8        | 5.6              | 9.6              |
|                  | Malignant             | 9120.485 | 0.002 | 0.347 | 1085  | 76.5        | 8.0              | 15.5             |
| Benign           | Malignant             | 9120.485 | 0.005 | 0.499 | 721   | 75.2        | 8.0              | 16.8             |
Fig. 1 Fitting Sloan (2007) near neutral model with the healthy tissue microbiome as source community and benign tumor tissue microbiome as destination community.

Fig. 2 Fitting Sloan (2007) near-neutral model with the healthy tissue microbiome as source.
Fig. 3 Fitting Sloan (2007) near-neutral model with the benign tissue microbiome as source community and malignant tumor tissue microbiome as destination community.

Fig 4. The percentage of neutral, below neutral (negatively selected) and above neutral (positively selected) species in the three types of metacommunity settings: “normal to benign”, “normal to malignant” and “benign to malignant”. The “benign to malignant” group shows the lowest neutral-species percentage but highest percentage of the above-neutral species.

Determining the neutrality at local community and metacommunity levels with Harris et al. (2017) MSN (multi-site neutral) model

We fitted the Harris et al.’s MSN model to four multi-site meta-community settings (groups). The first group consisted of three microbiomes from normal, benign tumor and malignant tumor...
tissues, respectively (Group name: “normal & benign & malignant”). The full results of fitting the MSN model to this group were listed in Table S1. Each of other three groups consisted of two microbiomes from two types of breast tissues. The group names of these three meta-communities were “normal & benign”, “normal & malignant” and “benign & malignant”. The full results of fitting to these three groups were listed in Table S2. Table 3 listed the mean parameters and corresponding standard errors of MSN models fitted to the breast microbiomes, which were summarized from Tables S1 and S2. Figure 5 shows an example of fitting the MSN model with the dataset of “normal & benign & malignant” group.

From these results, we found that all tested datasets passed the neutrality test with the MSN model (p-values > 0.9). It suggested that, at community level, the community assembly and diversity maintenance of breast tissue microbiome were driven predominantly by stochastic neutral forces, including dispersal and drift. The selections from tumor formation and malignant transformation were too weak to be detected. Table 3 also lists the average migration rates (M-values) between two types of breast tissue microbiomes. The microbial migration rates between normal and two types of tumor tissue were similar without significant difference (Wilcoxon test: p-value = 0.348 > 0.05), and the average M of them was 74.65. The migration rate between microbiomes of two tumor tissues was 54.321, and was significantly smaller than M-values of “normal & benign” and “normal & malignant” (Wilcoxon test: p-value < 0.001).

**Table 3.** The summary results of fitting Harris *et al’s* (2017) HDP-MSN (hierarchical Dirichlet process, multi-site neutral) model to the breast microbiomes, excerpted from Tables S1 and S2 in the OSI where the full fitting results from 1000 or 100 times of re-sampling were exhibited *

| Group                        | L₀      | θ       | M-value | Meta-community | Local community |
|------------------------------|---------|---------|---------|----------------|-----------------|
|                              | L_M    | N_M    | N       | p_M           | L_L            | N_L    | N     | p_L  |
| Normal & Benign & Malignant  | -5597.422 | 771.918 | 67.704  | -5130.309 | 2351 | 2500 | 0.940 | -5245.885 | 2317 | 2500 | 0.927 |
| Std. Err.                    | 68.235  | 11.424  | 0.824   | 58.384    | 9.012 | 0    | 0.004 | 61.189 | 6    | 0    | 0.002 |
| Normal & Benign              | -3833.236 | 637.265 | 73.711  | -3431.210 | 2410 | 2500 | 0.964 | -3545.737 | 2354 | 2500 | 0.942 |
| Std. Err.                    | 164.272 | 30.268  | 3.555   | 140.811  | 17.8  | 0    | 0.007 | 147.599 | 13.8 | 0    | 0.006 |
| Normal & Malignant           | -3500.268 | 544.794 | 75.589  | -3118.059 | 2399.1 | 2500 | 0.960 | -3220.030 | 2377.8 | 2500 | 0.951 |
| Std. Err.                    | 160.803 | 28.785  | 3.710   | 137.605  | 19.9  | 0    | 0.008 | 144.640 | 13.9 | 0    | 0.006 |
| Benign & Malignant           | -2825.108 | 501.805 | 54.321  | -2573.222 | 2336.9 | 2500 | 0.935 | -2651.725 | 2229.1 | 2500 | 0.892 |
| Std. Err.                    | 73.616  | 21.344  | 1.690   | 65.218   | 23.2  | 0    | 0.009 | 68.602 | 21.9 | 0    | 0.009 |

* N = 2500 is number of Gibbs samples selected from 25000 simulated communities (i.e., every tenth iteration of last 25000 Gibbs samples), chosen to compute pseudo p-value for conducting the neutrality test. L₀ is actual log-likelihood, computed from median of 25000 simulations and compared with log-likelihood of each simulated community. θ is median of biodiversity parameters computed from 25000 simulations. m is migration probability. M is average median of migration rates of local communities in each metacommunity (i.e., average median of individuals migrated per generation), computed from 25000 simulations. L_M is median of log-likelihoods of simulated neutral metacommunity samples. N_M is number of simulated neutral metacommunity samples with likelihoods exceeding L_M > L₀. P_M = N_M / N is pseudo p-value for testing neutrality at metacommunity level; if p_M > 0.05, metacommunity satisfies MSN model. L_L is median of log-likelihoods of simulated local community samples, and N_L is number of simulated local community samples with likelihoods exceeding L_L. P_L = N_L / N is pseudo p-value for testing neutrality at local community level; if p_L > 0.05, local community satisfies neutral model.

**Fig. 5 shows fitting of MSN to #2 sample by plotting predicted and observed species abundance rank distributions.**
Fig 5. An example illustrating the fitting of the MSN (multi-site neutral) model with the breast tissue microbiome: three samples, one from each of the three categories of samples (normal, benign and malignant), constitute a multi-site metacommunity.

Fig. 6. The box plot showing the fundamental dispersal number ($M$) in three metacommunity settings (groups): “normal & benign”, “normal & malignant” and “benign & malignant”. The rightmost group (benign & malignant) exhibited significantly smaller $M$-value than the two other groups (Wilcoxon test: $P$-value<0.001). There was no significant difference between the other two groups in their $M$-values ($P$-
value=0.348). Three standard summary numbers (statistics) of the parameter \( M \), including the first quartile (lower edge of the rectangle), median (the inside segment), third quartile (upper edge of the rectangle) were displayed, respectively. The “whiskers” above and below the box (rectangle) show the location of the minimum and maximum. The inter-quartile range (IQR) (showing the range of variation) is displayed by the height of the box; and the median shows the typical value. Outliers (<3xIQR or >3xIQR) are displayed outside the box.

**Determining the balance between niche selection and neutral drift with Tang and Zhou (2013) NNH (niche-neutral hybrid) model**

Similarly to the design of fitting the MSN model, we fitted the NNH model to four metacommunity groups: “normal & benign & malignant”, “normal & benign”, “normal & malignant”, and “benign & malignant”. The full results of fitting the NNH model to group of “normal & benign & malignant” were listed in Table S3, and to other three groups were listed in Table S4. Table 4 listed the average NNH parameters and corresponding standard errors, which are summarized from Table S3 and Table S4.

From these results, we found that there was no tested dataset passing the test with the NNH model \((p\text{-values} < 0.001)\). It further verified the previous finding, i.e., stochastic forces or neutral drift plays a predominant role in shaping the structure and diversity of breast tissue microbiome, while niche-differentiations or deterministic selection forces from tumor formation and malignant transformation plays little role.

**Table 4.** The summary results of fitting Tang & Zhou (2013) NNH (niche-neutral hybrid) model to the breast microbiomes, excerpted from Tables S3 and S4 in the OSI where the full fitting results from 1000 or 100 times of re-sampling were exhibited *

| Group                  | \( J \)   | \( S \)   | \( \theta \) | \( m \)   | \( x \)   | \( \gamma \) | \( R^2 \)  | \( \chi^2 \) | \( p\text{-value} \) | \( N_{\text{pass}} \) | \%(pass) |
|------------------------|-----------|-----------|------------|---------|---------|----------|-------|--------|----------------|---------------|---------|
| Normal & benign        | Mean      | 17022.2   | 267.730    | 79.143  | 0.000   | 0.840    | 0.966 | 0.423  | 236.880        | <0.001        | 0       |
|                        | Std. Err. | 1201.096  | 10.983     | 2.980   | 0.000   | 0.006    | 0.048 | 0.010  | 17.418         | 0.000         | 0       |
| Normal & malignant     | Mean      | 11605.9   | 252.500    | 75.978  | 0.000   | 0.838    | 0.980 | 0.389  | 215.120        | <0.001        | 0       |
|                        | Std. Err. | 935.399   | 15.140     | 3.733   | 0.000   | 0.006    | 0.047 | 0.010  | 17.776         | 0.000         | 0       |
| Benign & malignant     | Mean      | 12765.1   | 207.165    | 70.491  | 0.000   | 0.836    | 0.935 | 0.477  | 174.578        | <0.001        | 0       |
|                        | Std. Err. | 739.849   | 4.859      | 1.984   | 0.000   | 0.005    | 0.039 | 0.009  | 7.280          | 0.000         | 0       |
| Three groups           | Mean      | 13144.6   | 243.646    | 74.864  | 0.000   | 0.834    | 0.992 | 0.457  | 316.3          | <0.001        | 0       |
|                        | Std. Err. | 216.797   | 2.627      | 0.763   | 0.000   | 0.001    | 0.012 | 0.003  | 5.713          | 0.000         | 0       |

*J: average number of individuals per niche (local community) in each metacommunity, \( S \): average species number per niche (local community) in each metacommunity, \( \theta \): average fundamental biodiversity parameter per niche (local community) in each metacommunity, \( m \): average of migration coefficients, \( x \): average of birth to death ratio, \( \gamma \): average of migration rate, \( R^2 \): goodness-of-fit index, \( \chi^2 \): \( \chi^2 \)-value of chi-squared test for observed value against predicted value, \( p\text{-value} \) for \( \chi^2 \)-test; when \( p > 0.05 \), metacommunity satisfies NNH model. Last two columns are number and percentage of local communities (niches) that passed local neutrality test.

**Conclusions**

**The dispersal (migration) patterns of breast tissue microbiome**

In previous sections, based on Sloan near neural model (Table 2), we have demonstrated that while neutral species constitute, on average, approximately 85% species in the breast tissue microbiome, the selection effect from tumor progression does lead to certain percentage of
above-neutral or positively selected species. The percentage of positively selected species is about 10% on average, but could be up to about 17% in the progression to malignant tumor. The MSN/NNH models, nevertheless, showed that the selection is not sufficiently strong to lead to community/meta-community level dominance of stochastic neutral forces. In other words, whereas the breast tissue microbiomes are driven predominantly by stochastic neutral forces, at the species level, up to 17% approximately of species can be positively selected by tumor progression.

With MSN model, we can further infer the migration rate between breast tissue microbiomes. Fig 6 shows the fundamental dispersal number or the migration rate ($M$) of pair-wise microbial migration between two types of breast tissue microbiomes (normal tissue and benign tumor; normal and benign tumor, benign and malignant tumor). Average $M$-values were previously displayed in Table 3 and detailed $M$-values were displayed in Table S1 & S2. The non-parametric Wilcoxon tests revealed no significant difference in the $M$-value between “normal & benign” and “normal & malignant” ($p=0.348$), but the difference in $M$ was significant between “benign to malignant” and the two other groups mentioned above ($P$-value<0.001). This may indicate a significant difference between the benign and malignant tumor tissues in their microbial dispersal.

**Ethical Approval and Consent to participate**
Not applicable

**Consent for publication**
Not applicable

**Availability of supporting data**
All datasets analyzed in this study are available in public domain and download from Urbaniak et al. (2016)

**Competing interests**
The authors declare no competing interests.

**Funding**
This study received funding from the following sources: A National Natural Science Foundation of China (NSFC) Grant (No. 31970116) on Medical Ecology of Human Microbiome; The Cloud-Ridge Industry Technology Leader Award; An International Cooperation Grant (YNST) on Genomics & Metagenomics Big Data.

**Authors' contributions**
Lianwei Li: Software, Data Curation, Formal analysis, Writing-Original Draft.
Zhanshan (Sam) Ma: Conceptualization, Methodology, Writing-Reviewing and Editing.

**Acknowledgements**
We thank Dr. Camlllla Urbanlak and Dr. Gregor Reld for providing the datasets.
References

Burns AR, Stephens WZ, Stagaman K, et al. (2016) Contribution of neutral processes to microbial community assembly over host development. The ISME Journal, 10:655-664.

Etienne RS (2005) A new sampling formula for neutral biodiversity. Ecology Letters 8(3):253-260.

Etienne RS. (2007) A neutral sampling formula for multiple samples and an 'exact' test of neutrality. Ecology Letters 10(7): 608-618.

Goedert JJ, Jones G, Hua X, et al. (2015) Investigation of the Association Between the Fecal Microbiota and Breast Cancer in Postmenopausal Women: a Population-Based Case-Control Pilot Study. Journal of the National Cancer Institute, 107(8):djv147-djv147.

Harris K, Parsons TL, Ijaz UZ, et al (2017) Linking Statistical and Ecological Theory: Hubbell's Unified Neutral Theory of Biodiversity as a Hierarchical Dirichlet Process. Proceedings of the IEEE, 105(3):516-529.

Helmink BA, Khan MAW, Hermann A et al. (2019) The microbiome, cancer, and cancer therapy. Nature Medicine 25, 377–388.

Hieken TJ, Chen J, Hoskin TL, et al. (2016) The Microbiome of Aseptically Collected Human Breast Tissue in Benign and Malignant Disease. Scientific Reports, 6:30751.

Hubbell, SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, 2001.

Hubbell SP (2006) Neutral theory and the evolution of ecological equivalence. Ecology 87:1387–1398.

Hunt KM et al. (2011). Characterization of the diversity and temporal stability of bacterial communities in human milk. PLoS ONE 6, e21313.

Kwa M, Plottel CS, Blaser MJ et al. (2016) The Intestinal Microbiome and Estrogen Receptor-Positive Female Breast Cancer. Journal of the National Cancer Institute, 108(8).

Ma ZS, Guan Q, Ye CX, Zhang C, Foster J A, Forney LJ. (2015) Network analysis suggests a potentially ‘evil’ alliance of opportunistic pathogens inhibited by a cooperative network in human milk bacterial communities. Scientific Reports, 5: 8275.

Ma ZS, Li LW, Li WD, Li J, Chen HJ. (2016) Integrated network-diversity analyses suggest suppressive effect of Hodgkin’s lymphoma and slightly relieving effect of chemotherapy on human milk microbiome. Scientific Reports, 6: 28048.

McGill BJ. (2003). A test of the unified neutral theory of biodiversity. Nature, 422, 881–885.

Moossavi S, Sepehri S, Robertson B et al. (2019) Composition and Variation of the Human Milk Microbiota Are Influenced by Maternal and Early-Life Factors. Cell Host & Microbe, 25(2):324-335.

Nejman D, Livyatan I, Fuks G et al (2020) The human tumor microbiome is composed of tumor type-specific intracellular bacteria. 368: 973-980.
Poore G, Kopylova E, Zhu Q et al (2020) Microbiome analyses of blood and tissues suggest cancer diagnostic approach. 579: 567-574.

Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* 8:732–740.

Sloan WT, Woodcock S, Lunn M, Head IM, Curtis TP (2007) Modeling taxa-abundance distributions in microbial communities using environmental sequence data. *Microb Ecol* 53:443–455.

Tang J, Zhou S (2013) Hybrid niche-neutral models outperform an otherwise equivalent neutral model for fitting coral reef data. *Journal of Theoretical Biology*, 317(1):212-218.

Urbaniak C, Cummins J, Brackstone M, Macklaim JM, Gloor GB, Baban CK et al. (2014) Microbiota of human breast tissue. *Applied & Environmental Microbiology*. 80:3007–3014

Urbaniak C, Gloor GB, Brackstone M et al. (2016). The Microbiota of Breast Tissue and Its Association with Breast Cancer. *Applied & Environmental Microbiology*, 82(16):5039-5048.

Volkov I, Banavar JR, Hubbell SP, Maritan A (2003) Neutral theory and relative species abundance in ecology. *Nature* 424:1035–1037.

Volkov I, Banavar JR, Hubbell SP, Maritan A (2007) Patterns of relative species abundance in rainforests and coral reefs. *Nature* 450:45–49.