

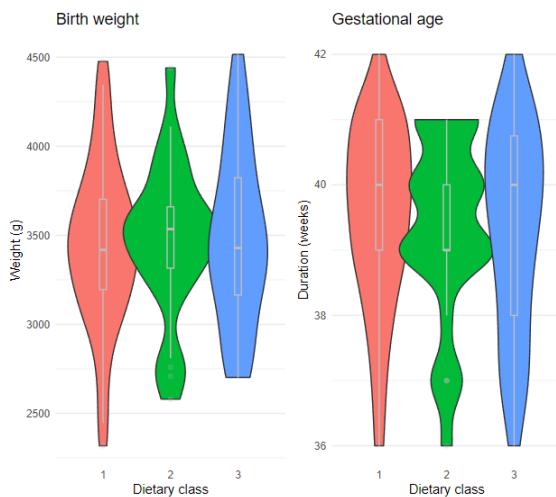
Supplementary figure 1



### Supplementary Figure 1. Descriptive statistics.

Proportions of infants in each dietary class for available categorical metadata. Results of Chi-square tests are indicated in the captions. Feeding data is provided for each time point.

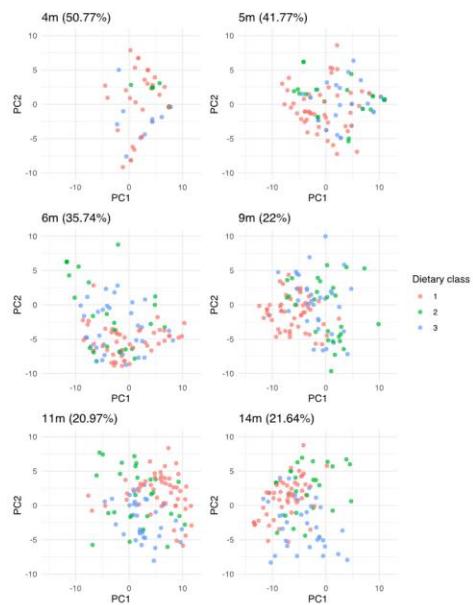
**Supplementary figure 2**



**Supplementary Figure 2. Descriptive statistics.**

Boxplots with the distribution of available numerical metadata additionally plotted. Colors represent dietary classes.

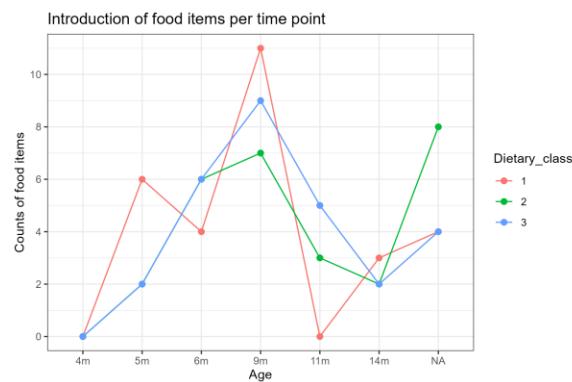
**Supplementary figure 3**



**Supplementary Figure 3. PCA score plots.**

The plots are representing first two PCs of logistic PCA models for infant dietary data at each time point, colored by dietary class. Each point represents an infant's position in the PC subspace, and the explained deviance is indicated in parentheses.

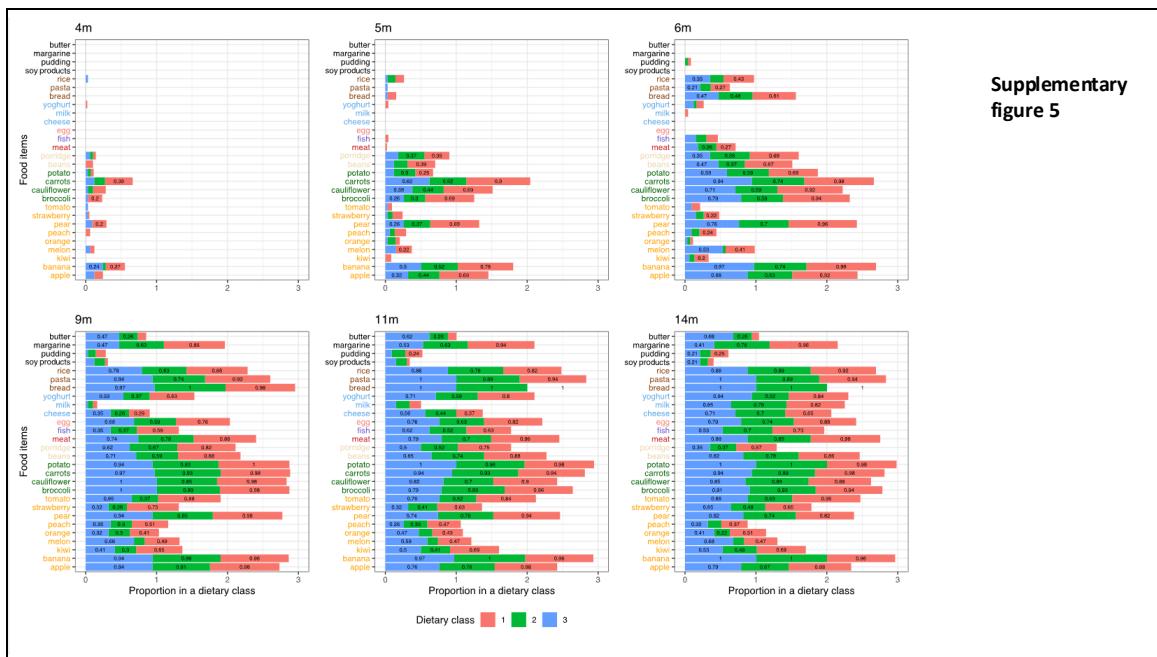
Supplementary figure 4



**Supplementary Figure 4. Food item introduction per time point for each dietary cluster.**

Food items introduced at the NA time point have never been received by more than 50% of infants within a dietary class.

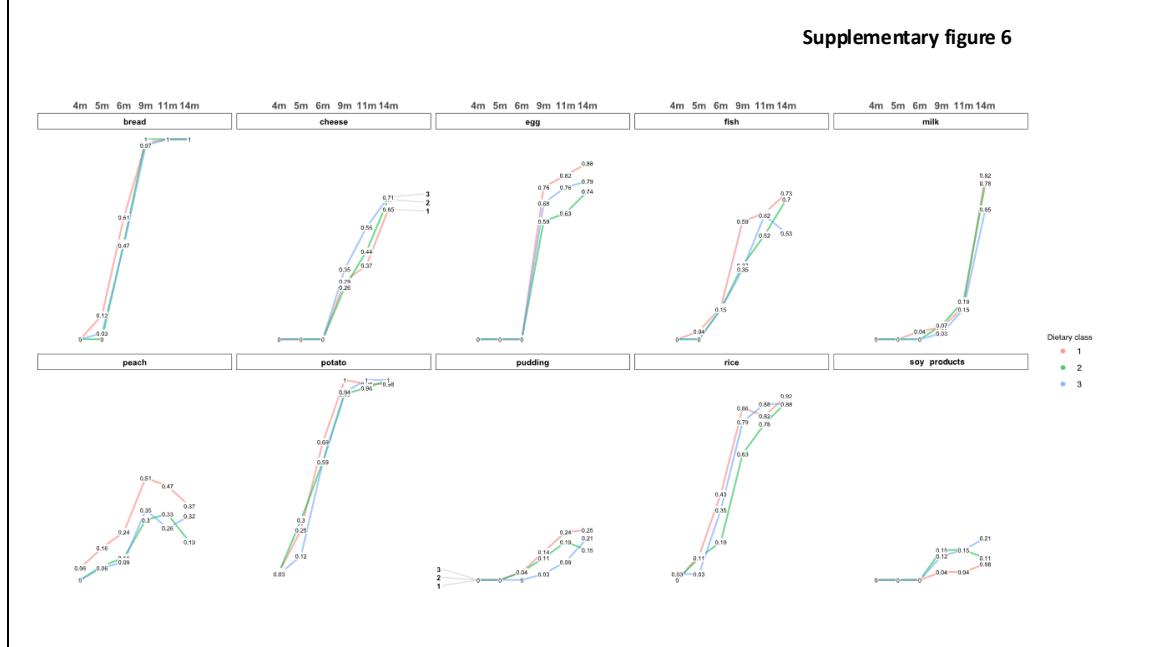
Supplementary  
figure 5



Supplementary Figure 5. Proportion of infants receiving food items over time, stratified by dietary class.

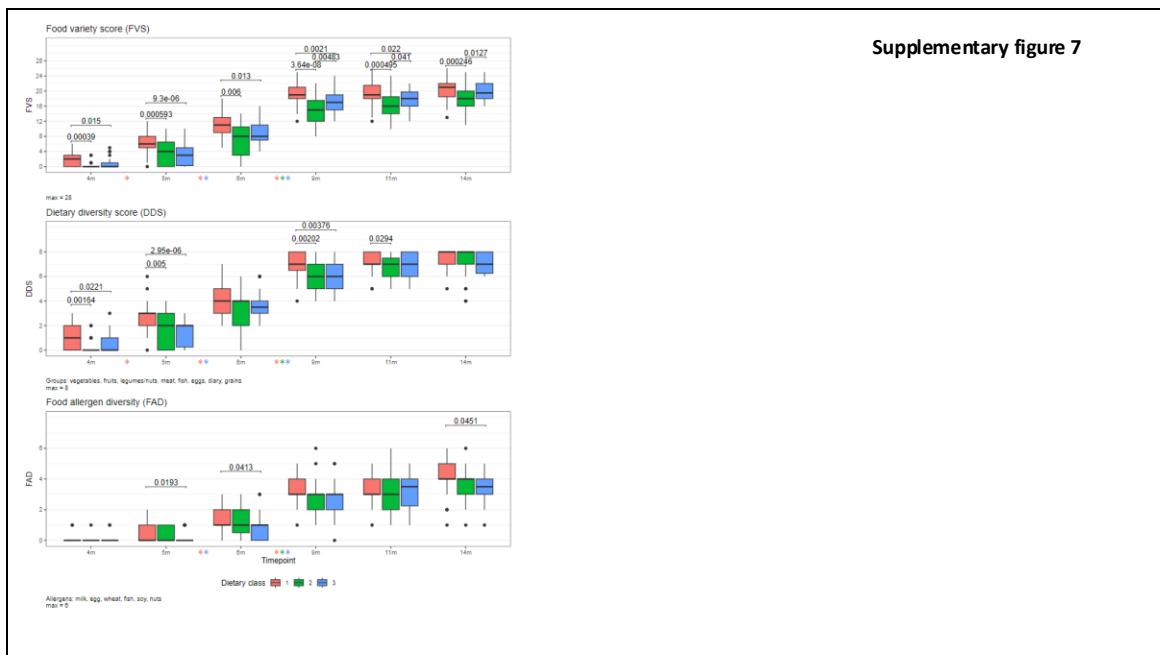
Bars for each food item were stacked and colored according to dietary class. Only proportion values exceeding 0.1 were additionally displayed.

Supplementary figure 6



Supplementary Figure 6. Changes in proportions of infants for non-significant food items (n=10) over time.

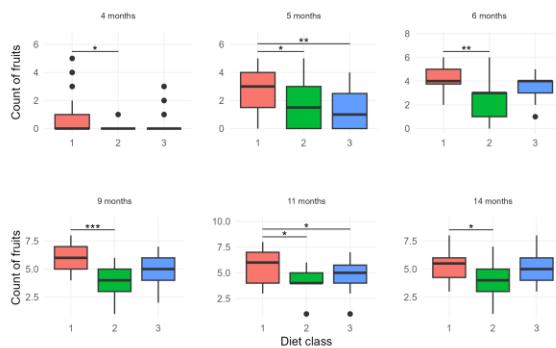
Supplementary figure 7



### Supplementary Figure 7. Boxplots of food diversity metrics for infants from each dietary class.

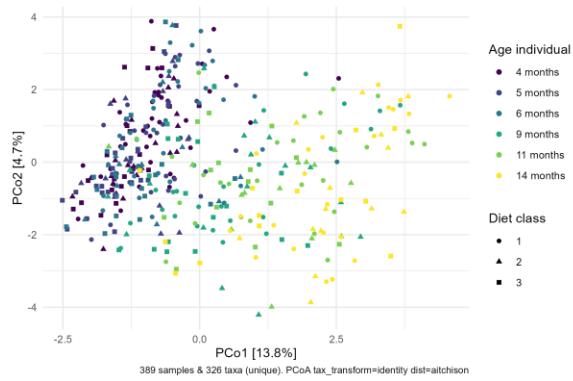
(A) The number of received food items by an infant at a specific time point. (B) The number of different dietary groups consumed by an infant at a specific time point. (C) The number of food allergens received by an infant at a specific time point. Statistically significant differences between dietary classes were displayed as brackets with adjusted p-values. Additionally, statistical differences between consecutive time points for each class were marked with "\*" below the x-axis (adjusted p-value).

Supplementary figure 8



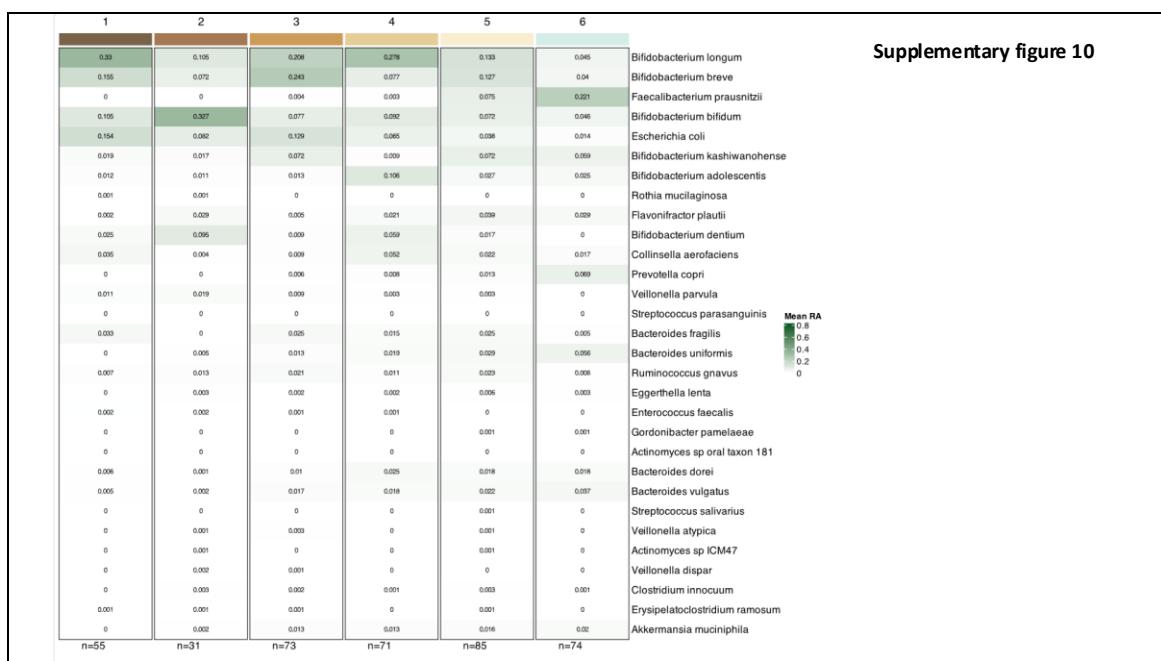
**Supplementary Figure 8. Boxplots of consumed fruits (max = 8) at each time point.**  
The adjusted p-values are represented with stars, where “\*\*\*” indicates p-values < 0.001, “\*\*” indicates p-values between 0.001 and 0.01, and “\*” indicates p-values between 0.01 and 0.05.

Supplementary figure 9



**Supplementary Figure 9. Principal Coordinate Analysis (PCoA) of fecal samples.**

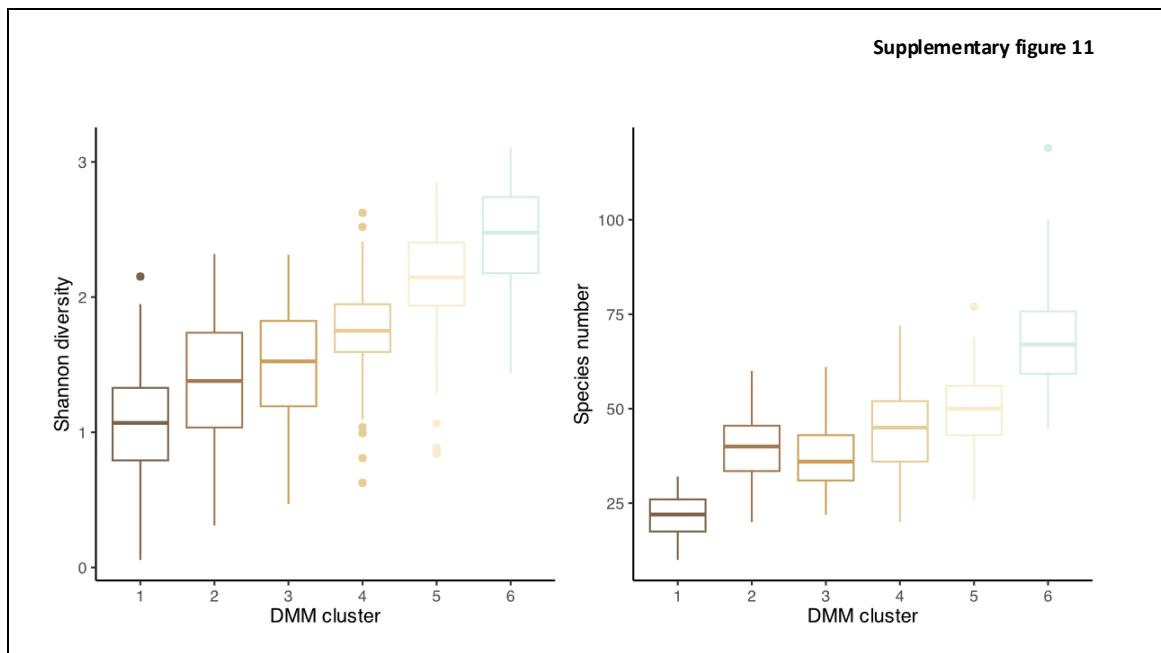
The samples are colored according to age, and shaped according to diet class.



### Supplementary Figure 10. Heatmap of bacterial abundance.

The heatmap displays the mean relative abundance of the 30 most dominant bacterial species for each DMM cluster.

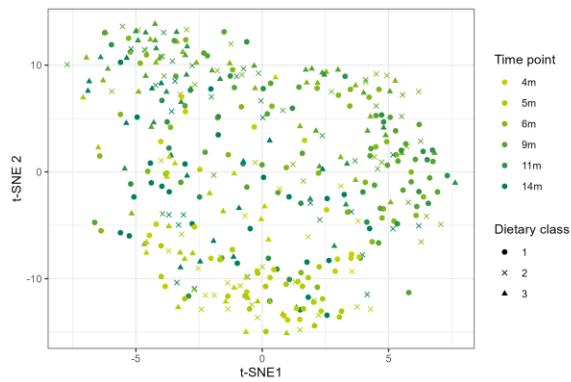
Supplementary figure 11



**Supplementary Figure 11. Box plots displaying diversity metrics for each DMM cluster.**

Shannon index and observed number of species are shown.

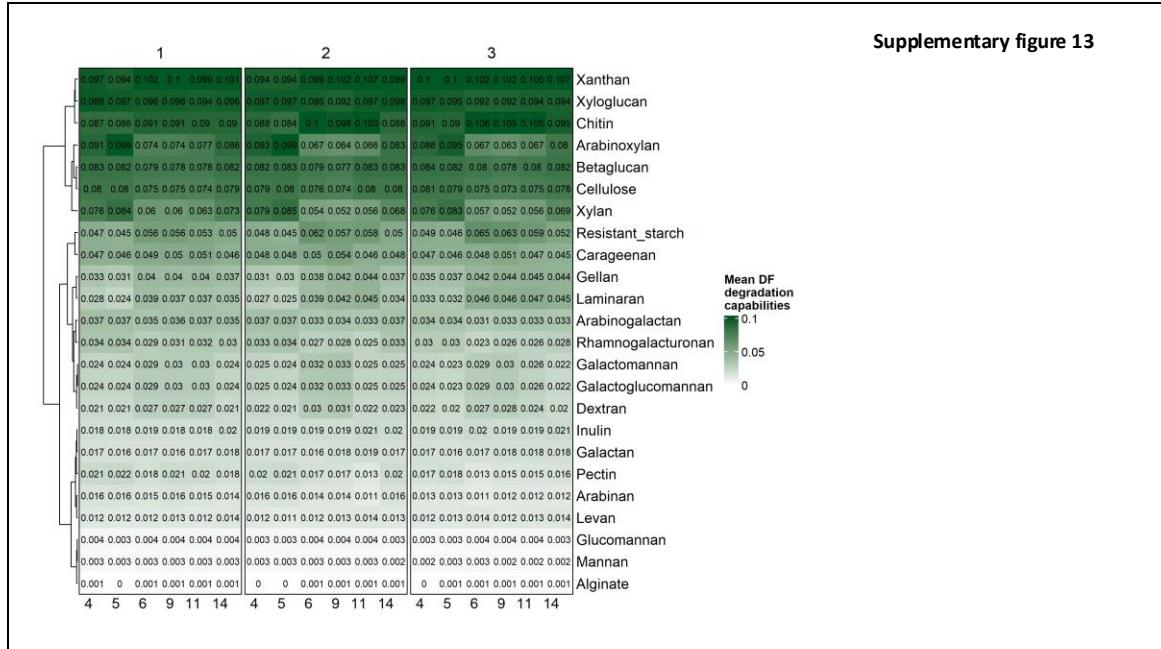
Supplementary figure 12



**Supplementary Figure 12. t-SNE dimensionality reduction plot.**

The first two dimensions of IFDPs t-SNE dimensionality reduction. IFDPs show variation over time rather than over dietary classes.

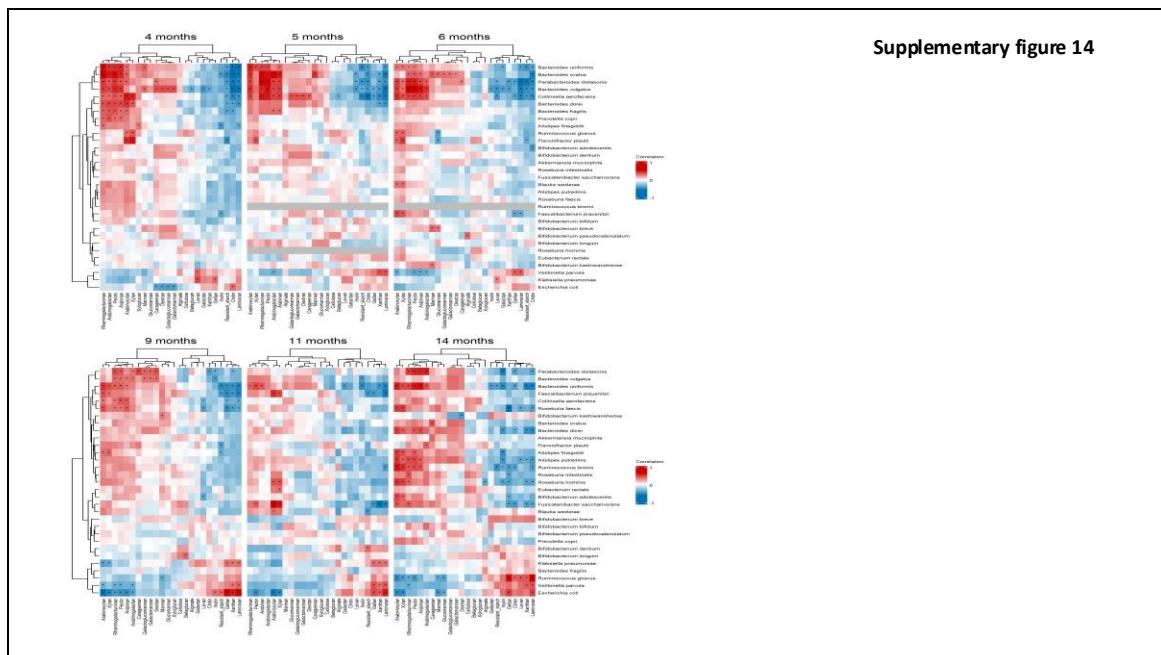
Supplementary figure 13



### Supplementary Figure 13. Heatmap of fiber abundance.

The heatmap displays mean degradation capabilities of each dietary fiber over time and by diet class.

Supplementary figure 14



**Supplementary Figure 14. Heatmap of Spearman correlations between dietary fiber and bacterial species.**

Bacterial species with a presence of at least 1% at each time point were included.

Significant correlations, with p-values adjusted using the Benjamin-Hochberg correction ( $\leq 0.05$ ), are marked with asterisks.