Statistical Model Let us use $Y_{i,j,k,l,m}(t)$ to denote the growth of the mth replicate/observational unit ($m = 1, 2$) at time t for the ith type, jth media, kth experiment, lth random isolate. We assume that the model is

$$\log\{Y_{i,j,k,l,m}(t)\} = \mu_{i,j,k,l,m}(t) + e_{i,j,k,l,m},$$

where $e_{i,j,k,l,m}$ are assumed to be iid Normal$(0, \sigma_e^2)$, and $\mu_{i,j,k,l,m}(t)$ represents the mean log-transformed growth at time t. Without a loss of generality we may assume that time t is in the compact interval $[0, 1]$. Now, with an appropriate choice of basis functions $\mu_{i,j,k,l}(t)$ can be approximated by

$$\mu_{i,j,k,l}(t) \approx \sum_{r=1}^{R} B_r(t)\psi_{r(i,j,k,l)},$$

where $B_1(t), \ldots, B_R(t)$ are R basis functions, and $\{\psi_{r(i,j,k,l)}, r = 1, \ldots, R\}$ are the regression coefficients. Specifically, in our calculation, B's are cubic B-splines with four inner knots at $(0.2, 0.4, 0.6, 0.6)$ and the boundary knots $(0, 1)$ when the time is scaled into $[0, 1]$.

Next we model the regression coefficients using the ANOVA type structure,

$$\psi_{r(i,j,k,l)} = \alpha_r + \beta_{r,i} + \gamma_{r,j} + \delta_{r,i,j} + a_{r,k,l}, \tag{1}$$

for $i = 1, 2$ (1: erm51N, 2: WTN), $j = 1, 2, 3$ (1: BHI, 2: MM, 3: MMNoFe), $k = 1, 2, 3$ (1: Experiment 1, 2: Experiment 2, 3: Experiment 3), $l = 1, \ldots, 6$ (1: Isolate 39, ..., 6: Isolate 77). Here β and γ's are the main effect of Type (strain) and media, while δ's are the interaction effect of Type (strain) and media. Since both types erm51N and WTN appear in each media, instead of a nested model/design, we consider a crossed model/design to analyze the data. For identifiability, we set $\beta_{r,1} = 0$ and $\gamma_{r,1} = 0$, $\delta_{r,1,j} = 0$ for $j = 1, 2, 3$, $\delta_{r,i,1} = 0$ for $i = 1, 2$ for every r. Here $a_{r,k,l}$'s are the mean zero random effect due to the kth experiment, and lth isolate. Specifically, we concatenate three random factors, and create this single random factor, call it a. To simply the computation and without no or negligible loss of information in most of practical cases, here we assume that the random effect of all R coefficients are the same, i.e., $a_{1,k,l} = \cdots = a_{R,k,l} = a_{k,l}$ (say), and assume that $a_{k,l}$ follows iid Normal$(0, \sigma_a^2)$.

Suppose that we are interested in testing if there is any effect of interaction effect between type (strain) and media on the growth. That means under the null hypothesis H_0, the interaction effect is zero. In terms of the parameters of model (1), $H_0 : \delta_{r,2,2} = \delta_{r,2,3} = 0$ for $r = 1, \ldots, R$ and the alternative hypothesis is H_a : at least one parameter of the set $\{\delta_{r,2,2}, \delta_{r,2,3}, r = 1, \ldots, R\}$ is non-zero. We shall use a likelihood ratio test

to carry out this test.

The number of basis functions R is the sum of the degree of the polynomial plus the number of inner knots plus one. In most of the cases, a cubic spline is good for approximating the underlying curve, and the number of inner knots could be around 3 to 5. Now, plugging the model of the coefficients in the spline model of Y, we have

$$\log\{Y_{i,j,k,l,m}(t)\} = \sum_{r=1}^{R} B_r(t)(\alpha_r + \beta_{r,i} + \gamma_{r,j} + \delta_{r,i,j}) + a_{k,l} \sum_{r=1}^{R} B_r(t) + e_{i,j,k,l,m}.$$

Since B-splines with intercept satisfies $\sum_{r=1}^{R} B_r(t) = 1$ for any $t \in [0,1]$, the above model reduces to

$$\log\{Y_{i,j,k,l,m}(t)\} = \sum_{r=1}^{R} B_r(t)(\alpha_r + \beta_{r,i} + \gamma_{r,j} + \delta_{r,k}) + a_{k,l} + e_{i,j,k,l,m}. \tag{2}$$

Model (2) involves with unknown fixed effect parameters $(\alpha_r, \beta_{r,i}, \gamma_{r,j}, \delta_{r,i,j}), r = 1, \ldots, R, i = 1, 2, j = 1, 2, 3,$ and the variance components σ_a^2 and σ_e^2.

For a given media, first the difference between the two growth curves for two strains was measured via the relative change of areas under the two curves. This measure is easy to interpret. However, a more appropriate measure to verify the difference between two curves is the integrated squared differences $\Delta = \int\{(\mu_{1,j,k,l}(t) - \mu_{2,j,k,l}(t)\}^2 dt$, where the random effects are all set to zero. Thus,

$$\Delta = \int \{\sum_{r=1}^{R} B_r(t)(\beta_{r,2} + \delta_{r,2,j})\}^2 dt.$$

Although it is easy to estimate Δ, numerical enumeration of the asymptotic null distribution of Δ is difficult. Therefore, we have used a parametric bootstrap procedure to estimate the null distribution (when $\beta_{r,2} = 0$ and $\delta_{r,2,j} =$ for all $j = 1, 2, 3$ and all r) of Δ. This helps to estimate the p-value.

2