Metagenomic Predictions of Growth and Carcass Traits in Pigs with the Use of Bayesian Alphabet and Machine Learning Methods.

C. Maltecca¹, D. Lu¹, F. Tiezzi¹, C. Schillebeeckx², N.P. McNulty², C. Schwab³ and C. Shull³
¹North Carolina State University, Raleigh, USA, ²Matatu Inc., Saint Louis, MO, USA, ³The Maschhoffs LLC., Carlyle, IL, USA.

Summary
In this paper, we evaluated the power of metagenome measures taken at three time points over the growth test period (weaning, 15 weeks and 22 weeks) to predict growth and carcass traits in a line of crossbred pigs. Models from the Bayesian alphabet (Bayesian Lasso) as well as two machine learning approaches (Random Forest and Gradient Boosting) were employed to predict weight, backfat, loin depth and loin area at week 15 and 22. Prediction accuracy was measured as correlation between true and predicted phenotypes in cross validation. In addition, a time dependent recurrent neural network using all microbiome measures simultaneously was fitted to classify individuals in 4 groups based on daily gain and backfat at week 22. In most cases prediction accuracy increased with the inclusion of microbiome composition. Accuracy was larger with the inclusion of microbiome composition taken at week 15 and 22, with values ranging from ~.30 for loin traits to > .50 for backfat. Model choice only affected prediction accuracy marginally. Microbiome can be used as an effective tool to predict growth and carcass in swine.

Keywords: metagenomic predictions, growth, swine, machine learning

Introduction
Efficiency of producing saleable meat products is largely determined by costs associated with feed and by the amount of and quality of lean meat produced (Hoque et al. 2009). Utilizing feed resources more efficiently has become a clear challenge that faces the livestock industry. Recent efforts have been devoted to identify and exploit the genomic variability of individual pigs in increasing feed efficiency (Jiao et al. 2014b), (Jiao et al. 2014a), (Howard et al. 2015), (Lu et al. 2017). Nonetheless, a continued effort concentrating only on the pig variability for efficiency will inevitably result in diminished marginal gains, incurring in concomitant losses of overall fitness and diversity over time (Colleau and Tribout 2008). The amount and type of bacteria present in the gut of individuals represent a key part of all mammalian organisms (Gill et al. 2006). The makeup of the microbiome represents a vast pool of genomic diversity that contributes to the individual physiology and health (Pfluighoef and Versalovic 2012). Particularly, the intestinal microbiome directly affects the degradation of carbohydrates, provides short chain fatty acids, mitigates and alter the effect of potential toxic compounds and produce essential vitamins (Gill et al. 2006). Different composition of the gut population in humans has been linked to the ability of degrading enzymes, maintain a certain population balance, influence the overall health status and control fatness and growth. Though gut microbial diversity in pigs has been described to some extent, composition and function of a healthy microbial ecosystem however, have yet to be qualitatively and quantitatively defined and used as a tool to maximize animal health and performance. Within this paper, we assessed the power of metagenomic predictions based on fecal samples, to foresee growth and carcass composition in a population of crossbred healthy pigs. In doing so we employed machinery typical of host genomics predictions, including models of the Bayesian alphabet, as well as non-parametric machine learning algorithms.

Materials and methods
Data. From a Duroc closed-nucleus population 28 boars were selected to be sires of the
individuals used in this study. Sires were mated to crossbred sows to generate terminal-cross piglets. These were weaned at an average of 19 days of age and grouped in single-sire-gender pens (groups). During the nursery, growth and finish period, all pigs were fed standard diets. End of test was declared on a pen-specific basis, entire pens of pigs were taken off test and sent for harvest at a pen mean live weight of 304.6 ± 5.51 lb. Live weight measurements were taken individually at the start (weaning) and end of the study as well as at weeks 15, 18 and 22 post-weaning. Rectal swabs were collected from all pigs in a pen at 3 time points, including weaning (wean), 15 weeks post weaning (15wk), and 22 weeks post weaning (22wk). Four pigs were chosen randomly per pen carcass growth measurements, and their rectal swabs were used for microbiome sequencing. In the end, the number of samples at weaning, 15 weeks, and 22 weeks were 1205, 1295, and 1283, respectively. There were 1039 animals having samples collected at all 3 time points.

Bacterial genomic DNA (gDNA) was extracted from each rectal swab by mechanical disruption in phenol:chloroform. Phased, bi-directional amplification of the V4 region (515–806) of the 16S rRNA gene was employed to generate indexed libraries for Illumina sequencing using the strategy described by Faith et al. (2013). All sequencing was performed at the DNA Sequencing Innovation Lab at the Center for Genome Sciences and Systems Biology at Washington University in St. Louis. Pairs of V4 16S rRNA gene sequences were first merged into a single sequence using FLAsh v1.2.11 (Magoč and Salzberg 2011). Sequences with a mean quality score below Q35 were then filtered out using PRINSEQ v0.20.4 (Schmieder and Edwards 2011). Sequences with >97% nucleotide sequence identity were then clustered into OTUs. A modified version of GreenGenes (Schloss and Handelsman 2006) was used as the reference database. The most abundant sequence in each cluster was used as the representative sequence for the OTU. Sparse OTUs were filtered out by requiring a minimum total observation count of 1200 for an OTU to be retained, and the OTU table was rarefied to 10,000 counts per sample. A total of 380 OTUs with complete phylogeny assigned where retained for further analyses.

Phenotypes included in the analyses were: weight (wt), loin area (lea), loin depth (ld) and backfat (bf) at week 15 and 22. Additionally, average daily gain (adg) was obtained from wean to 22 weeks.

**Training and testing sets.** A stratified fivefold cross validation scheme was used to recursively randomly split data into training (70% of observations) and prediction (30% of observations) sets, maintaining equal representation of the 28 sires present in the trial.

**Regression analysis.** For this investigation, each combination of trait/time was treated as a separate analysis and accuracy of prediction was obtained as the average correlation between predicted and measured phenotypes in the test sets. **Bayesian Regression:** OTU counts per individual obtained as outlined above were used in a Bayesian Lasso regression model using the BGLR R package (Pérez and Campos 2014). For each phenotype, the model fitted included fixed effects of sex (2 levels), plus the effect of sire (28 levels), as well as a covariate for weaning weight. Pen was fitted as a random effect with a scaled inverted Chi-square prior distribution. OTU counts were centered and scaled and were fitted to the model with the use of a double exponential distribution (Bayesian Lasso, BL). The hyper parameter was fixed and its value was assigned through a grid search on the full dataset (results not shown). **Random Forest and Gradient Boosting:** Two of the most flexible and popular machine learning methods were chosen for the analysis, Gradient Boosting (GB) and Random Forest (RF). Both are ensemble methods for weak predictors. Each model fits many decision trees aimed at improving the predictive accuracy, while controlling over-fitting. Gradient Boosting builds an additive model in a forward-wise fashion. The algorithm is then optimized
based on a loss function. For the current data, the deviance loss function was employed. Conversely RF fits several decision trees on various sub-samples of the dataset. The quality of split is than measured through different criteria. For the current analysis mean square error (MSE) was employed. Model performance for both methods is dependent on a series of parameters governing the behavior of the models. These parameters where again set through a grid search on the whole dataset (data bot shown). Both approaches were modeled using the scikit-learn package in python 3.6 (Pedregosa et al. 2011). Predictors included in the models were the same as for the BL models.

**Classification analysis. Recurrent Neural Network.** The analyses in the previous section do not account for the temporal dependence of microbiome samples within individuals. To exploit this dependence, we employed a Recurrent Neural Network (RNN) model, in a classification setting. RNN are models naturally suited to accommodate temporal sequenced data. Within this paper, we employed a gated recurrent unit model (GRU) (Chung et al. 2014), to classify individuals into one of 4 classes obtained through a k-means partitioning. Based on adg wean to 22 weeks and bf at week 22 the following classes were established: high growth high fat (HgHf), low growth low fat (LgLf), low growth high fat (LgHf) and high growth low fat (HgLf). All OTU counts at each time point were in this case fitted simultaneously. The RNN was implemented with the use of the TensorFlow API for python 3.6 ("TensorFlow White Papers" 2017). Accuracy of classification was in this case measured with the use of a receiver operator characteristic curve (ROC).

**Results**

Mean and standard deviations for the traits used in the regression analysis are reported in table 1.

<table>
<thead>
<tr>
<th></th>
<th>wt</th>
<th>bf</th>
<th>lea</th>
<th>ld</th>
<th>adg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Me</td>
<td>151.8</td>
<td>258.8</td>
<td>0.490</td>
<td>0.78</td>
<td>4.39</td>
</tr>
<tr>
<td>an</td>
<td>21.82</td>
<td>29.52</td>
<td>11</td>
<td>0.20</td>
<td>0.71</td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Regression:* In tables 1 to 3 are reported accuracies of prediction for each trait fecal microbiome timepoint combination. Microbiome contribution to prediction was measured as deviation from a null model which included only effects of sex, sire weight at weaning and pen. OTUs inclusion in the prediction models increased accuracies in most instances with respect to the null model. Nonetheless, the amount varied according to the microbiome timepoint. In general, inclusion of microbiome composition at weaning had low predictive power for traits measured at week 15 and 22, with the null model correlations in the test set ranging from a minumum of 0.06 (ld 15wk) to a max of 0.37 (bf 22wk) and substatially identical to the ones including the microbiome, regardless of the model employed. Only for the combination BL/bf 22wk the model including OTUs was better than the null one (0.41) albeit with a larger SD.

**Table 2. Cross validation correlation (SD) for growth and carcass traits and OTU counts at weaning.**

*Null Model included only fixed effects of sex and sire covariate of weight at weaning and random effect of pen.

<table>
<thead>
<tr>
<th></th>
<th>wt</th>
<th>bf</th>
<th>lea</th>
<th>ld</th>
<th>adg</th>
</tr>
</thead>
<tbody>
<tr>
<td>wean</td>
<td>15</td>
<td>22</td>
<td>15</td>
<td>22</td>
<td>15</td>
</tr>
<tr>
<td>Null</td>
<td>0.09</td>
<td>0.11</td>
<td>0.25</td>
<td>0.37</td>
<td>0.16</td>
</tr>
</tbody>
</table>
Microbiome composition at week 15 increased substantially the accuracy in the test sets. The amount was dependent on the trait/time combination. In general, and as expected, microbiome composition increased prediction accuracies more for traits measured concomitantly with the microbiome sampling. For wt at week 15 inclusion of OTUs resulted in an accuracy (averaged across models) of 0.39, with an increase of 0.27 over the null model. Prediction accuracy for wt week 22 was roughly half, still the inclusion of OTUs boosted accuracy of .063 (averaged across models) with respect to the null model. Similar trends were seen in bf week15 where microbiome composition lifted accuracy to .494 (averaged across models), .23 more than the null model.

Table 3. Cross validation correlation for growth and carcass traits and OTU counts at week 15.

<table>
<thead>
<tr>
<th>Model</th>
<th>15wk</th>
<th>wt</th>
<th>bf</th>
<th>lea</th>
<th>ld</th>
<th>adg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null*</td>
<td>0.12</td>
<td>0.26</td>
<td>0.41</td>
<td>0.19</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>BL</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>GB</td>
<td>0.11</td>
<td>0.16</td>
<td>0.18</td>
<td>0.36</td>
<td>0.11</td>
<td>0.15</td>
</tr>
<tr>
<td>RF</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Similar trends were seen with the inclusion of microbiome measured at week 22. Again, back fat was the trait benefitting the most from the inclusion of microbiome data with an increase of ~.10 (averaged across models) compared to the null model. As for the previous time points lea and ld benefitted the least from OTUs inclusion. It should be noted that given the temporal succession of sampling, combinations of traits at week 15 and microbiome at week 22 were not fitted.

Table 4. Cross validation correlation for growth and carcass traits and OTU counts at week 22.

<table>
<thead>
<tr>
<th>Model</th>
<th>22wk</th>
<th>wt</th>
<th>bf</th>
<th>lea</th>
<th>ld</th>
<th>adg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null*</td>
<td>0.42</td>
<td>0.23</td>
<td>0.50</td>
<td>0.44</td>
<td>0.34</td>
<td>0.34</td>
</tr>
<tr>
<td>BL</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>GB</td>
<td>0.40</td>
<td>0.25</td>
<td>0.47</td>
<td>0.44</td>
<td>0.34</td>
<td>0.34</td>
</tr>
<tr>
<td>RF</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Model choice was in all cases a wash, with performances from BL, GB and RF substantially overlapping, and without a clear winner across combinations of trait/time. Bayesian Lasso models appeared nonetheless to be slightly less sensitive to the splitting of the data.
**Classification:** In figure 1 are reported the ROC curves for the classification analysis. The dotted line represents the average ROC curve (averaged across 4 classes), while each solid line represents the ROC for an individual class. Different panels depict results from a different fold in the cross validation. The average AUC was 0.632. In all folds and as expected, best predictions were achieved for the HgHf, LgLf lables (AUC of ~0.70 in both cases). The worse performance was obtained for LgHf, with AUC close to the .50 line representing the chance. Interestingly the ability of predicting individuals with HgLf (the ones more suitable for population improvement) was intermediate (AUC ~0.60).

*Figure 1. Receiver operator curve for classification of growth. Each panel represent a fold.*

**Conclusions**

Microbiome composition can be effectively used as a predictor of growth and composition traits, particularly for fatness traits. Inclusion of OTU predictors could potentially be used to promote fast growth of individuals while limiting fat accumulation. Early microbiome measures might not be good predictors of growth and OTU information might be best collected at later life stages. Future research should focus on the inclusion of both microbiome ad well as host genome information in predictions, as well as the interaction between the two. Furthermore, the influence of microbiome on feed efficiency as well as carcass and meat quality should be investigated.

**Acknowledgments**

Funding for the current work is partially provided by the National Pork Board (2016-1314), the North Carolina Pork Council (2017-1905, 2017-1929), NCSU, Matatu and the
List of References


