Published in final edited form as:

J Pediatr. 2015 February ; 166(2): 282–288.e5. doi:10.1016/j.jpeds.2014.10.065.

Computational Gene Expression Modeling Identifies Salivary Biomarker Analysis that Predict Oral Feeding Readiness in the Newborn

Jill L. Maron, MD, MPH^{1,*}, Jooyeon S. Hwang¹, Subash Pathak, MS², Robin Ruthazer, MPH², Ruby L. Russell³, and Gil Alterovitz, PhD³

¹Mother Infant Research Institute at Tufts Medical Center, Division of Newborn Medicine, Floating Hospital for Children at Tufts Medical Center, Boston, MA

²Research Design Center/Biostatistics Research Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA and Tufts Clinical and Translational Science Institute, Tufts University, Boston, MA

³Center for Biomedical Informatics, Harvard Medical School, Boston, MA

Abstract

Objective—To combine mathematical modeling of salivary gene expression microarray data and systems biology annotation with RT-qPCR amplification to identify (phase I) and validate (phase II) salivary biomarker analysis for the prediction of oral feeding readiness in preterm infants.

Study design—Comparative whole transcriptome microarray analysis from 12 preterm newborns pre- and post-oral feeding success was used for computational modeling and systems biology analysis to identify potential salivary transcripts associated with oral feeding success (phase I). Selected gene expression biomarkers (15 from computational modeling; 6 evidence-based; and 3 reference) were evaluated by RT-qPCR amplification on 400 salivary samples from successful (n=200) and unsuccessful (n=200) oral feeders (phase II). Genes, alone and in combination, were evaluated by a multivariate analysis controlling for sex and post-conceptual age (PCA) to determine the probability that newborns achieved successful oral feeding.

Results—Advancing post-conceptual age ($p < 0.001$) and female sex ($p = 0.05$) positively predicted an infant's ability to feed orally. A combination of five genes, *NPY2R* (hunger signaling), *AMPK* (energy homeostasis), *PLXNA1* (olfactory neurogenesis), *NPHP4* (visual behavior) and *WNT3* (facial development), in addition to PCA and sex, demonstrated good accuracy for determining feeding success (AUROC = 0.78).

© 2014 Elsevier Inc. All rights reserved.

***Corresponding Author:** Jill L. Maron MD, MPH, Associate Professor of Pediatrics, Mother Infant Research Institute, Tufts Medical Center, Division of Newborn Medicine, jmaron@tuftsmedicalcenter.org, 617-636-0766 (phone), 617-636-1469 (fax).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The authors declare no conflicts of interest.

Conclusions—We have identified objective and biologically relevant salivary biomarkers that noninvasively assess a newborn’s developing brain, sensory and facial development as they relate to oral feeding success. Understanding the mechanisms that underlie the development of oral feeding readiness through translational and computational methods may improve clinical decision making while decreasing morbidities and health care costs.

Preterm births affect an estimated 11.5% of all pregnancies in the United States resulting in medical costs exceeding \$26 billion annually¹. Prior to discharge, each infant must demonstrate mature oral feeding skills in accordance to the American Academy of Pediatrics’ guidelines². The determination of oral feeding readiness in the preterm newborn remains a significant clinical challenge³. Oral feeding is a complex developmental task requiring maturation and integration of the nervous, gastrointestinal, sensory, skeletal muscular and hypothalamic systems⁴. Disruption or delayed maturation in one or several of these developmental systems may result in choking, feeding aversion, and poor growth⁵. Further, infants either born at term gestation or who correct to term post-conceptual age (PCA) who cannot successfully orally feed are at increased risk for developmental disabilities^{6–7}. Due to the biological complexities of oral feeding, caregivers have been limited to subjective feeding assessment tools or “best guess” clinical assessments to determine the feeding readiness of preterm newborns^{8–10}. This, in turn, has resulted in significant feeding associated morbidities, prolonged length of stay, and millions of dollars in health care expenditure. A recent Cochrane Review assessing the benefits of neonatal feeding assessment tools concluded that “there is no evidence to inform clinical practice”, highlighting the strong need for novel approaches to assess oral feeding readiness in the preterm newborn¹¹.

Transcriptomic analysis of neonatal salivary samples represents an innovative and noninvasive strategy to monitor, in real-time, the gene expression patterns of the multiple biological and developmental systems required for oral feeding readiness¹². In this study, we combined computational modeling of gene expression microarray data and *a priori* systems biology knowledge with highthroughput reverse-transcription quantitative polymerase chain reaction (RT-qPCR) amplification to identify and validate objective and biologically relevant salivary biomarkers predictive of neonatal oral feeding readiness.

Methods

This study was approved by the Tufts Medical Center Institutional Review Board, with parental consent. Both preterm and term neonates (gestational age ≥ 37 weeks) were recruited for this study. For the majority of enrolled subjects, PCA was based upon dating by first trimester ultrasound. In the rare instant when a first trimester assessment was not available, second trimester imaging was used to determine the age of the infant. Feeding status of infants was determined with the use of a cue based feeding assessment tool¹³. Infants ≥ 32 weeks’ PCA were allowed to feed if they maintained a stable cardio-respiratory status, demonstrated appropriate feeding cues and tolerated enteral nutrition. Percent oral feeding success was calculated by dividing the volume of enteral nutrition taken orally by the total volume of enteral nutrition provided in the day. Successful oral feeders took 100% of their feeds by mouth; unsuccessful oral feeders took < 100% of feeds orally. A chi-

squared test was performed between successful and unsuccessful oral feeders to assess the possibility that human derived breast milk was impacting gene expression.

Salivary samples were collected with techniques developed in our laboratory and previously described¹⁴. Saliva was sampled approximately one hour after a feed to limit contamination with breast milk or formula. Samples were only collected during the day to reduce potential effects of circadian rhythms on gene expression. Saliva was immediately stabilized with 500 μ L of RNAProtect saliva (Qiagen, Venio Limburg, Netherlands), vortexed, and placed at 4°C for a minimum of 48 hours prior to total RNA extraction with Qiagen's RNeasy Mini Kit (Venio Limburg, Netherlands) per manufacturer's instructions. On column DNase digestion occurred for each sample. Extracted total RNA (14 μ L) was stored at -80°C awaiting downstream analysis.

Phase I: Biomarker Discovery

Two salivary samples were collected from preterm infants: one pre- and one post-oral feeding success. Total RNA was amplified with the NuGEN™ Ovation Pico WTA system (San Carlos, CA) per manufacturer's instructions. Five micrograms of amplified RNA were fragmented and biotinylated with the NuGEN™ Encore Biotin module prior to hybridization onto Affymetrix™ HG U133a Plus 2.0 arrays (Santa Clara, California). Each array was washed and stained in the GeneChip® Fluidics Station 400, scanned with the GeneArray Scanner, and analyzed using the GeneChip Microarray Suite 5.0 (Affymetrix™).

Identification of Biomarkers—Gene expression data were normalized with GenePattern prior to statistical analyses. Robust Multi-Array Average and quantile normalization were applied to microarray data. Statistical analyses used Multi-Experiment Viewer. Statistical significance was set at $p < 0.01$. Genes that met statistical criteria and were differentially expressed pre- and post-oral feeding success were enriched via Database for Annotation, Visualization and Integrated Discovery (DAVID) to further explore their biological function. DAVID is a publically available database that provides researchers with comprehensive functional annotation tools to better understand the biological relationships of lists of genes (<http://david.abcc.ncifcrf.gov/home.jsp>)^{15–16}. Three genes were cross-listed in the Bayes network: *chromosome 4 open reading frame 34 (C4orf34)*, *olfactory receptor, family 5, subfamily AK, member 4 pseudogene (OR5AK4P)*, *myosin regulatory light chain interacting protein (MYLIP)*. Using 10 fold cross validation, an area under the receiver operator characteristic curve (AUROC) of 0.90 for the Bayesian network with 200 best features was calculated.

Phase II: Biomarker Validation

Development of the RT-qPCR Platform—All efforts were made to adhere to the minimum information of publication of quantitative real-time PCR experiments (MIQE) guidelines¹⁷. Genes were only considered for the platform if they were identified within the Bayesian Network or were shown previously to be associated with feeding success^{12, 18–19}. Each potential gene candidate had a systems biology review with Ingenuity Pathway Analysis® and PubMed.org to determine its potential role in oral feeding. Candidate genes were incorporated onto the platform if they were involved in one or more of the following

developmental systems: sensory integration, hypothalamic regulation of feeding, facial development, neurodevelopment and/or gastrointestinal development. Due to budget limitations, the platform was restricted to 24 genes, inclusive of three reference genes. Pertinent information regarding all genes on the platform is listed in Table I (available at www.jpeds.com).

Extracted RNA was converted to cDNA with the SuperScript VILO cDNA Synthesis kit (Life Technologies) per manufacturer's instructions. Next, we performed a selective pre-amplification of all gene targets on the platform with a pooled custom assay mix prepared by Life Technologies. Amplified samples were diluted 1:5 with RNase-free water prior to qPCR. A positive control was run on all plates. Median Ct values and confidence intervals for the positive control reference genes were calculated to assess plate-to-plate variability. A negative control was run once to ensure that no genes amplified in the absence of substrate. The thermal-cycling profile was as follows: UNG incubation at 50°C for 2 minutes, polymerase activation at 95°C for 20 seconds, and 40 cycles of PCR with denaturation at 95°C for one second and annealing and extension at 60°C for 20 seconds. Only samples with successful amplification of all reference genes (beta actin [*ACTB*], glyceraldehyde 3-phosphate dehydrogenase [*GAPDH*], and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta [*YWHAZ*]) were included in the analysis. All genes were run in duplicate. In the event of discrepant results between duplicates (e.g., one well with gene amplification, one well without gene amplification), gene expression was considered positive. It took three days to acquire, process and analyze each saliva sample on the RT-qPCR platform and cost approximately \$50/sample. A flow chart depicting the study methods can be found in Figure 1 (available at www.jpeds.com).

Statistical Analyses

All analyses were performed using the SAS system for Windows, version 9.3 (SAS Institute Inc., Cary, NC). Unadjusted associations between gene expression, dichotomized as expressed or not expressed with feeding status, dichotomized as feeder or non-feeder, were evaluated using chi-square tests with the salivary sample used as the unit of analyses. Sensitivity, specificity, positive predictive value, and negative predictive value for select genes were computed. Logistic regression was used to screen genes, sex, and PCA for inclusion in a multivariable model using a stepwise selection process with a threshold p-value of 0.10 set for both entering and staying in the model. Genes selected using the selection process, PCA, and sex were used in a generalized estimating equation (GEE) analysis with a generalized linear model using a logit link function to get adjusted odds ratios and 95% confidence intervals for the association with the outcome of feeding while accounting for the multiple samples from some infants. The variable estimates from this model were used to generate estimated probabilities of feeding for each sample and a corresponding ROC area. Grids of estimated probabilities of feeding for different combinations of genes, PCA, and sex were constructed. As a sensitivity analysis, a proportional odds model was used to model the association of the feeding gradient (categorized into 5 levels of oral feeding success: 1=0–20%, 2=21–40%, 3=41–60%, 4=61–80%, 5=81–100%) using the same variables that were included in the multivariable GEE model. The assumption of proportional odds was tested using a score test.

Results

Phase I-Biomarker Discovery

Twelve neonatal subjects provided salivary samples for phase I. Infants' PCA ranged from 33 5/7 to 36 2/7 weeks pre-oral feeding success and 34 2/7 to 37 3/7 weeks post-oral feeding success.

All quality measurements were within recommended guidelines for Affymetrix™ arrays as described by Gentleman et al²⁰. Control probes showed similar data quality between experiments.

Four hundred and twenty one genes met statistical significance (paired t-test df=11, p<0.01). Following enrichment with DAVID, 141 genes were incorporated into a Bayesian Network to predict feeding status. Of these genes, 15 were incorporated onto the RT-qPCR platform. The remaining six genes were evidence based^{12, 18–19}.

Phase II-Biomarker Validation

There were 298 infants who provided salivary samples for biomarker validation. Of these, 70 infants provided more than one saliva sample at a different PCA and feeding stage. Saliva volumes ranged from 10–20 µL.

There were 361/400 (90%) samples that successfully amplified all reference genes. These samples were obtained from 269 of the original 298 subjects. Their ages ranged from 31 4/7 to 44 5/7 weeks' PCA for unsuccessful oral feeders (n=167), and 32 5/7 to 48 2/7 weeks' PCA for successful oral feeders (n=194). The gene *NK6 homeobox 3 (NKX6-3)* failed to amplify in any sample. There was no gene amplification in our negative control wells. Quality controls for positive control reference genes were as follows: *ACTB*: median Ct 29.38, SD 0.83, CV 2.81; *YWHAZ*: median Ct 30.52, SD 0.72, CV 2.35; *GAPDH*: median Ct 32.69, SD 17.6, CV 97. There was 3.6% discordance in gene amplification in duplicate wells. There was no statistically significant difference between infants who received breast milk in either group (77% of successful oral feeders vs. 69% of unsuccessful oral feeders, p = 0.069).

The percentage of successful oral feeders was compared between samples with and without gene expression, between males and females, and between older and younger infants. These data and the corresponding unadjusted p values are shown in Table II (available at www.jpeds.com). A stepwise selection procedure identified the most informative genes on the platform (p < 0.10). Advancing PCA, modeled as a continuous variable, was highly significant in our analysis (p<0.001). Further, females were more likely to successfully feed orally than males across all PCAs (p=0.05). After controlling for PCA and sex, five genes were identified: *Plexin A1 (PLXNA1)*, *Neuropeptide Y2 receptor (NPY2R)*, *AMP-activated protein kinase (AMPK)*, *wingless-type MMV integration site family, member 3 (WNT3)*, and *nephronophthisis 4 (NPHP4)*. Unadjusted odds ratios for feeding for each gene in our final multivariable model, together with sensitivity, specificity, positive and negative predictive values are shown in Table III. Adjusted odds ratios of each gene and corresponding p-values for feeding after controlling for PCA, sex and each other can be found in Table IV. These

genes were combined into a model to predict successful oral feeders. The AUROC was 0.78 (Figure 2). Measurement estimates from multivariable models of feeding are in Table V (available at www.jpeds.com).

A test of the proportional odds assumption was not significant ($p = 0.22$). Only *NPY2R* could discriminate feeding status across quintiles of feeding success ($p = 0.03$). The odds ratio estimates for each gene used in the ordinal logistic regression model can be found in Table IV.

Fitted probability tables were generated using gene expression for combinations of genes in the multivariable GEE model, PCA and sex. These predictive tables provided the fitted likelihood an infant could sustain full oral feeds. An example of these models for males and females at 35 weeks' PCA can be found in Figure 3. The predictive tables for PCAs ranging from 33 to 40 weeks' PCA can be found in Figure 4 (available at www.jpeds.com).

Discussion

Subjective assessment and long-standing nursing protocols dictate feeding regimens in the NICU. These limited, qualitative assessment tools have not been properly validated and fail to identify the biological mechanisms responsible for failed oral feeding attempts¹¹. By combining computational modeling of gene expression microarray data with systems biology knowledge we identified biologically relevant salivary biomarkers that provide a noninvasive window into a newborn's developing brain, sensory systems, and facial development as they relate to oral feeding readiness.

The five most predictive genes of feeding success represent a diverse range of biological systems including sensory integration (*NPHP4*, *PLXNA1*), hypothalamic regulation of feeding (*NPY2R*), facial development (*WNT3*), and energy expenditure (*AMPK*). Each gene was informative in a binary fashion. A mature oral feeding pattern was predicted when three genes demonstrated negative gene expression (*NPHP4*, *NPY2R*, and *WNT3*), and two genes had positive gene expression (*AMPK*, *PLXNA1*). We hypothesized that no one gene could explain the biological complexities of oral feeding maturation in the newborn. Thus, we included genes with an adjusted p value up to < 0.10 in our predictive modeling. Combined, our panel of five biomarkers had good accuracy for determining feeding success (AUROC = 0.78), similar to the accuracy of newly developed scoring systems to detect prostate cancer²¹ and far exceeding common genetic variants to predict breast cancer²². Further, predictive models combining salivary gene expression profiles, PCA and sex, not only assessed the probability of feeding success, but most importantly highlighted specific developmental pathways that were likely contributing to feeding immaturity. This finding is unique to our assay. Ultimately, these models have the potential to allow caregivers to develop individualized treatment modalities based upon each infant's gene expression profile to improve quality of care.

Our laboratory demonstrated the important role of hypothalamic hunger signaling for successful neonatal oral feeding. Previously, we reported that detection of *NPY2R*, a well described hunger modulator^{23–24} in neonatal saliva, was highly predictive of oral feeding

immaturity¹⁹. Down regulation of *NPY2R* gene expression is associated with obesity and excessive weight gain in animal models²⁵. In our prior study, we showed that decreased expression of *NPY2R* likely contributed to the physiological hyperphagia required for rapid weight gain in a neonate's first year of life¹⁹. Here, we have validated our prior findings by demonstrating that infants who could successfully orally feed were less likely to express *NPY2R* compared with those who could not. Further, of the five biomarkers identified in this study, only *NPY2R* could significantly discriminate level of feeding success. These data strongly suggest that *NPY2R* is an informative oral feeding biomarker.

Similarly, *AMPK* is a well described regulator of whole-body energy balance. It is a serine/threonine kinase that is ubiquitously expressed in all organ systems²⁶. Activation of *AMPK* in the hypothalamus increases feeding and body weight gain²⁷. We demonstrated that infants with detectable levels of *AMPK* had three times the odds of being a successful oral feeder compared with infants without detectable levels of *AMPK* in saliva. Interestingly, *AMPK* directly interacts with neuropeptide Y neurons in the hypothalamus to regulate feeding in animal models^{28–29}. Although further studies are needed to delineate the precise role that *AMPK* plays in neonatal oral feeding, the gene nevertheless appears to be differentially expressed between those infants who can and who cannot successfully orally feed.

An infant must also be able to demonstrate a coordinated suck, swallow, and breathing pattern in order to safely feed by mouth. We included genes on the platform involved in facial development to evaluate if ongoing facial morphogenesis affected a preterm infant's oral feeding readiness. For example, *WNT3* is a member of the Wnt family of secreted glycoproteins and is important for lip and palate development^{30–31}. Mutations within the gene are associated with cleft lip and/or palate³². Our study suggests that down-regulation of this gene is associated with feeding success. We speculate that lack of expression of *WNT3* correlates with mature facial development, which in turn allows for an integrated suck, swallow, and breathing pattern.

Maturation of sensory integration pathways is also required for oral feeding success³³. Here, we identified two sensory genes, *PLXNA1* and *NPHP4*, that appear to influence neonatal oral feeding. *PLXNA1* is a semaphorin receptor that controls axon guidance, including axonal growth of olfactory sensory neurons³⁴. McIntyre et al showed that expression of *PLXNA1* is increased in mature compared with developing olfactory sensory neurons³⁵. We found that successful oral feeders were significantly more likely to have detectable levels of *PLXNA1* in their saliva compared with unsuccessful oral feeders, suggesting that the gene may serve as an indicator of olfactory maturation. Finally, *NPHP4* is a cilia-associated protein required for normal photoreceptor ribbon synapse maintenance and outer retinal segment formation³⁶. Similar to *PLXNA1*, we hypothesize that this gene is involved in mature sensory development and is an integral part of the multisensory integration known to be required for oral feeding success. Verhagen *et al.* and others have shown that food perception and 'reach-to-eat' movements in human infants are dependent upon the integration of the visual, audition and somatosensation systems^{37–38}.

A limitation of this study is that an infant's individual feeding status was determined, in part, by subjective nursing assessment; this may have skewed the data. This limitation highlights the fact that there is currently no objective assay to assess neonatal feeding readiness and there is a strong need to identify relevant biomarkers that can accurately establish when a preterm infant has developed sufficient maturation in some, if not all, of these pathways. In addition, this study represents the practices of only a single NICU. Future multi-center trials must be performed to determine the applicability of salivary gene expression analysis to predict feeding readiness in the preterm infant across institutions.

We were also unable to delineate the precise biological mechanism(s) by which these genes affect feeding behavior. Rather, we have shown the biological plausibility that each gene is related to oral feeding and demonstrated a strong association with a gene's expression profile and feeding status. Nevertheless, gene expression analysis of these five biomarkers has the potential to significantly improve current feeding practices in the NICU.

In conclusion, we have identified objective and developmentally relevant salivary biomarkers associated with feeding success in the newborn. Predictive data modeling provides assessments of a mature feeding pattern when taking into account an infant's PCA and sex. Compared with subjective cue-based feeding algorithms, this approach represents a significant advance to supplement clinical decision making and establishes the foundation for the development of a point-of-care assay to provide caregivers with an objective assessment of neonatal developmental feeding readiness to reduce feeding related morbidities and associated health care costs.

Acknowledgments

We would like to thank the families who graciously participated in this study, as well as the staff of the Neonatal Intensive Care Unit and Mother Infant Unit at the Floating Hospital for Children at Tufts Medical Center, Boston, MA

Funded by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (K08 HD 059819-05), the Gerber Foundation, and the Richard B. Saltonstall Charitable Fund (xx [to J.M.] and NIH UL1TR001064 [o the Tufts Clinical and Translational Science Institute]).

Abbreviations and Acronyms

AMPK	<i>AMP-activated protein kinase</i>
NPHP4	<i>nephronophthisis 4</i>
NPY2R	<i>Neuropeptide Y2 receptor</i>
PLXNA1	<i>Plexin A1</i>
WNT3	<i>wingless-type MMV integration site family, member 3</i>
PCA	post-conceptual age
RT-qPCR	reverse-transcription quantitative polymerase chain reaction

References

1. March of Dimes [www.marchofdimes.org]. 2013 Preterm Birth Report. Available from : <http://www.marchofdimes.com/mission/prematurity-reportcard.aspx#>.
2. Committee on Fetus and Newborn, Hospital discharge of the high-risk neonate. *Pediatrics*. 2008; 122:1119–1126. [PubMed: 18977994]
3. Lau C, Alagurusamy R, Schanler RJ, Smith EO, Shulman RJ. Characterization of the developmental stages of sucking in preterm infants during bottle feeding. *Acta Paediatr*. 2000; 89:846–852. [PubMed: 10943969]
4. Barlow SM. Central pattern generation involved in oral and respiratory control for feeding in the term infant. *Curr Opin Otolaryngol Head Neck Surg*. 2009; 17:1871–1893.
5. Amaizu N, Shulman RJ, Schanler RJ, Lau C. Maturation of oral feeding skills in preterm infants. *Acta Paediatr*. 2008; 97:61–67. [PubMed: 18052999]
6. Mizuno K, Ueda A. Neonatal feeding performance as a predictor of neurodevelopmental outcome at 18 months. *Dev Med Child Neurol*. 2005; 47:299–304. [PubMed: 15892371]
7. Samara M, Johnson S, Lamberts K, Marlow N, Wolke D. Eating problems at age 6 years in a whole population sample of extremely preterm children. *Dev Med Child Neurol*. 2010; 52:e16–e22. [PubMed: 19832883]
8. Palmer MM, Crawley K, Blanco IA. Neonatal oral-motor assessment scale: a reliability study. *J Perinatol*. 1993; 13:28–35. [PubMed: 8445444]
9. Howe T-H, Lin K-C, Fu C-P, Su C-T, Hsieh C-L. A review of psychometric properties of feeding assessment tools used in neonates. *J Obstet Gynecol Neonatal Nurs*. 2008; 37:338–349.
10. Bingham PM, Ashikaga T, Abbasi S. Relationship of neonatal oral motor assessment scale to feeding performance of preterm infants. *J Neonatal Nurs*. 2012; 18:30–36. [PubMed: 22639541]
11. Crowe L, Chang A, Wallace K. Instruments for assessing readiness to commence suck feeds in preterm infants: effects on time to establish full oral feeding and duration of hospitalization. *Cochrane Database Syst Rev*. 2012; 4:CD005586. [PubMed: 22513933]
12. Maron JL. Insights into neonatal oral feeding through the salivary transcriptome. *Int J Pediatr*. 2012; 195:153
13. Ludwig SM, Waitzman KA. Changing feeding documentation to reflect infant-driven feeding practice. *Newborn Infant Nurs Rev*. 2007; 7:155–160.
14. Dietz JA, Johnson KL, Wick HC, Bianchi DW, Maron JL. Optimal techniques for mRNA extraction from neonatal salivary supernatant. *Neonatology*. 2012; 101:55–60. [PubMed: 21791940]
15. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc*. 2009; 4:44–57. [PubMed: 19131956]
16. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009; 37:1–13. [PubMed: 19033363]
17. Bustin SA, Benes V, Garson JA, et al. The MIQE guidelines: Minimum information of publication of quantitative real-time PCR experiments. *Clin Chem*. 2009; 55:611–622. [PubMed: 19246619]
18. Maron JL, Johnson KL, Rocke DM, Cohen MG, Liley AJ, Bianchi DW. Neonatal salivary analysis reveals global developmental gene expression changes in the preterm infant. *Clin Chem*. 2009; 55:184–191.
19. Maron JL, Johnson KL, Dietz JA, Chen ML, Bianchi DW. Neuropeptide Y2 receptor (NPY2R) expression in saliva predicts feeding immaturity in the preterm neonate. *PLoS One*. 2012; 7:e37870. [PubMed: 22629465]
20. Gentleman, R.; Carey, VJ.; Huber, W.; Irizarry, RA.; Dudoit, S. *Bioinformatics and computational biology solutions using R and Bioconductor*. New York: Springer; 2005.
21. Ma W, Diep K, Fritsche HA, Shore N, Albitar M. Diagnostic and prognostic scoring system for prostate cancer using urine and plasma biomarkers. *Genet Test Mol Biomarkers*. 2014; 18:156–163. [PubMed: 24512523]

22. Hüsing A, Canzian F, Beckmann L, Garcia-Closas M, Diver WR, Thun MJ, et al. Prediction of breast cancer risk by genetic risk factors, overall and by hormone receptor status. *J Med Genet.* 2012; 49:601–608. [PubMed: 22972951]
23. Lin S, Boey D, Herzog H. NPY and Y receptors: lessons from transgenic and knockout models. *Neuropeptides.* 2004; 38:189–200. [PubMed: 15337371]
24. Butler AA, Cone RD. Knockout models resulting in the development of obesity. *Trends Gene.* 2001; 17:S50–S54.
25. Naveillhan P, Hassain H, Canals JM, Ekstrand AJ, Larefalk A, Chhajlani V, et al. Normal feeding behavior, body weight and leptin response require the neuropeptide Y Y2 receptor. *Nat Med.* 1999; 5:1188–1193. [PubMed: 10502824]
26. Ronnett GV, Ramamurthy S, Kleman AM, Landree LE, Aja S. AMPK in the brain: its roles in energy balance and neuroprotection. *J Neurochem.* 2009; 109:17–23. [PubMed: 19393004]
27. Lage R, Diéguez C, Vidao-Puig A, López M. AMPK: a metabolic gauge regulating whole-body energy homeostasis. *Trends Mol Med.* 2008; 14:539–549. [PubMed: 18977694]
28. Kohno D, Sone H, Minokoshi Y, Yada T. Ghrelin raises $[Ca^{2+}]_i$ via AMPK in hypothalamic arcuate nucleus NPY neurons. *Biochem Biophys Res Commun.* 2008; 366:388–392. [PubMed: 18068666]
29. Kohno D, Sone H, Tanaka S, Kurita H, Gantulga D, Yada T. AMP-activated protein kinase activates neuropeptide Y neurons in the hypothalamic arcuate nucleus to increase food intake in rats. *Neurosci Lett.* 2011; 499:194–198. [PubMed: 21658429]
30. Soeren L, Ganner A, Walz G. Inversin, Wnt signaling and primary cilia. *Differentiation.* 2012; 83:S49–S55.
31. Jiang R, Bush JO, Lidral AC. Development of the upper lip: Morphogenetic and molecular mechanisms. *Dev Dyn.* 2006; 235:1152–1167. [PubMed: 16292776]
32. Menezes R, Letra A, Kim AH, Küchler EC, Day A, Tannure PN, et al. Studies with *Wnt* genes and nonsyndromic cleft lip and palate. *Birth Defects Res A Clin Mol Teratol.* 2010; 88:995–1000. [PubMed: 20890934]
33. Schaal B, Coureaud G, Doucet S, Delaunay-El Allam M, Moncomble AS, Montigny D, et al. Mammary olfactory signalisation in females and odor processing in neonates: Ways evolved by rabbits and humans. *Beh Brain Res.* 2009; 200:346–358.
34. Winberg ML, Noordermeer JN, Tamagnone L, Comoglio PM, Spriggs MK, Tessler-Lavigne M, Goodman CS Plexin A is a neuronal semaphorin receptor that controls axon guidance. *Cell.* 1998; 95:903–916. [PubMed: 9875845]
35. McIntyre JC, Titlow WB, McClintock TS. Axon growth and guidance genes identify nascent, immature, and mature olfactory sensory neurons. *J Neurosci Res.* 2010; 88:3243–3256. [PubMed: 20882566]
36. Won J, De Evsikova CM, Smith RS, et al. NPHP4 is necessary for normal photoreceptor ribbon synapse maintenance and outer segment formation, and for sperm development. *Hum Mol Genet.* 2011; 120:482–496. [PubMed: 21078623]
37. Verhagen JV, Englelen L. The neurocognitive bases of human multimodal food perception: Sensory integration. *Brain Res Rev.* 2006; 30:316–650.
38. Sacrey L-AR, Karl JM, Whishaw IQ. Development of visual and somatosensory attention of the reach-to-eat movement in human infants aged 6 to 12 months. *Exp Brain Res.* 2012; 223:121–136. [PubMed: 22948738]

Figure 1.
Flow chart depicting the methods of this study.

Figure 2.

Area under the receiver operating characteristic curve (AUROC). The five most promising genes identified on our platform (*PLXNA1*, *NPY2R*, *AMPK*, *NPHP4*, *WNT3*) were combined to predict oral feeding success. Area under the curve = 0.78.

Figure 3.

Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when *NPY2R*, *WNT3* and *NPHP4* were not expressed (NEG/-), while *AMPK* and *PLXNA1* were expressed (POS/+)

Figure 4.

1. Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when NPY2R, WNT3 and NPHP4 were not expressed (NEG/-), while AMPK and PLXNA1 were expressed (POS/+).

2. Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when NPY2R, WNT3 and NPHP4 were not expressed (NEG/-), while AMPK and PLXNA1 were expressed (POS/+).
3. Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when NPY2R, WNT3 and NPHP4 were not expressed (NEG/-), while AMPK and PLXNA1 were expressed (POS/+).
4. Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when NPY2R, WNT3 and NPHP4 were not expressed (NEG/-), while AMPK and PLXNA1 were expressed (POS/+).
5. Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when NPY2R, WNT3 and NPHP4 were not expressed (NEG/-), while AMPK and PLXNA1 were expressed (POS/+).
6. Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when NPY2R, WNT3 and NPHP4 were not expressed (NEG/-), while AMPK and PLXNA1 were expressed (POS/+).
7. Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when NPY2R, WNT3 and NPHP4 were not expressed (NEG/-), while AMPK and PLXNA1 were expressed (POS/+).
8. Fitted probability of feeding success based upon PCA, sex and gene expression profiles. Probability of feeding success was favorable when NPY2R, WNT3 and NPHP4 were not expressed (NEG/-), while AMPK and PLXNA1 were expressed (POS/+).

Table 1

Pertinent information regarding the genes included on the RT-qPCR platform.

System	Gene	Function	Life Technologies ID	Spans Exons	Contains SNPs	Amplicon Length
Hypothalamic Regulation of Feeding Behavior	<i>NPY2R</i>	Feeding behavior	Hs01921296	No	No	143
	<i>AMPK</i>	Feeding behavior, energy expenditure	Hs01562308	Yes	No	93
Sensory Integration	<i>TCF7L2</i>	Feeding behavior	Hs01009044	Yes	No	89
	<i>PLXNA1</i>	Olfactory system development	Hs00413698	Yes	No	53
	<i>OR5AK4P</i>	Olfactory system development	Hs03454034	Yes	No	119
	<i>GNAS</i>	Sensory perception of smell	Hs00255603	Yes	No	59
	<i>POU2F1</i>	Olfactory placode formation	Hs00231250	Yes	No	80
Neurodevelopment	<i>NHP4</i>	Visual behavior	Hs00296416	Yes		67
	<i>KCNJ10</i>	Visual perception	Hs01922935	No	No	101
	<i>NRP1</i>	Axon guidance	Hs00826128	Yes	No	90
	<i>NNAT</i>	Brain development	Hs00193590	Yes	No	72
	<i>NPAS3</i>	Social development	Hs00223201	Yes	No	88
Facial Development	<i>IL1RAPL2</i>	Central nervous system development	Hs00213600	Yes	No	90
	<i>ONECUT2</i>	Peripheral nervous system development	Hs00191477	No	No	57
	<i>NKX6-3</i>	Central nervous system development	Hs00332012	Yes	No	113
	<i>WNT3</i>	Cell fate and patterning	Hs00229135	Yes	No	69
	<i>RUNX1</i>	Skeletal development	Hs00231079	Yes	No	64
Gastrointestinal Development	<i>MAP2K</i>	Pharyngeal development	Hs00983247	Yes	No	68
	<i>TSHZ3</i>	Embryonic development	Hs02379784	No	No	115
	<i>ABCB1</i>	Hydrolase activity	Hs00184491	Yes	No	110
	<i>INSR</i>	Utilization and uptake of glucose	Hs00961554	Yes	No	68
	<i>GAPDH</i>	Reference Gene	Hs99999905	Yes	Yes	124
Controls	<i>YWHAZ</i>	Reference Gene	Hs03044281	Yes	No	106
	<i>ACTB</i>	Reference Gene	Hs01060665	Yes	No	63

Table 2

Unadjusted comparisons of feeding status between sample subgroups

Subgroup Strata	% (Ratio) of Successful Oral feeders		p-value *
<i>Gene</i>	<u>Infants with Positive Expression</u>	<u>Infants with Negative Expression</u>	
<i>ABCB1</i>	52.8% (57/108)	54.2% (137/253)	0.81
<i>ONECUT2</i>	38.5% (5/13)	54.3% (189/348)	0.26
<i>NNAT</i>	42.9% (27/63)	56.0% (167/298)	0.06
<i>IL1RAPL2</i>	37.5% (15/40)	55.8% (179/321)	0.03
<i>NPAS3</i>	52.4% (11/21)	53.8% (183/340)	0.90
<i>WNT3</i>	41.8% (33/79)	57.1% (161/282)	0.02
<i>RUNX1</i>	54.5% (183/335)	42.3% (11/26)	0.22
<i>POU2F1</i>	55.2% (186/337)	33.3% (8/24)	0.04
<i>GNAS</i>	53.7% (194/361)		
<i>NPHP4</i>	51.1% (113/221)	57.9% (81/140)	0.21
<i>NKX6-3</i>		53.7% (194/361)	
<i>PLXNA1</i>	56.1% (165/294)	43.3% (29/67)	0.06
<i>NRP1</i>	55.0% (181/329)	40.6% (13/32)	0.11
<i>INSR</i>	54.2% (181/334)	48.2% (13/27)	0.54
<i>MAP2K1</i>	53.9% (194/360)	0% (0/1)	0.28
<i>TCF7L2</i>	55.0% (176/320)	43.9% (18/41)	0.18
<i>ACTB</i>	53.7% (194/361)		
<i>AMPK</i>	55.0% (187/340)	33.3% (7/21)	0.05
<i>NPY2R</i>	49.0% (76/155)	57.3% (118/206)	0.11
<i>KCNJ10</i>	54.9% (118/215)	52.1% (76/146)	0.60
<i>TSHZ3</i>	49.6% (121/244)	62.4% (73/117)	0.02
<i>YWHAZ</i>	53.7% (194/361)		
<i>OR5AK4P</i>	50.5% (96/190)	57.3% (98/171)	0.18
<i>GAPDH</i>		53.7% (194/361)	
Demographics	Male	Female	
<i>Sex</i>	48.4% (105/217)	69.8% (89/144)	0.01
-			
-	Younger (PCA < 36 weeks)	Older (PCA ≥ 36 weeks)	
<i>Age</i>	37.9% (64/169)	67.7% (130/192)	<0.01
-			

* p-values from Chi-square test

Table 3

Unadjusted associations of Genes and Demographics with Feeding outcome

Genes	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Odds Ratio	Odds Ratio 95% CI	p value
<i>PLXNA1</i>	85.05	22.75	56.12	56.72	2.89	(1.47, 5.67)	0.002
<i>AMPK</i>	96.36	8.38	55	66.67	3.21	(1.09, 9.48)	0.03
<i>WNT3</i>	17.01	72.46	41.77	42.91	0.59	(0.33, 1.07)	0.09
<i>NPY2R</i>	39.18	52.69	49.03	42.72	0.71	(0.36, 1.0)	0.05
<i>NPHP4</i>	58.25	35.33	51.13	42.14	0.6	(0.34, 1.03)	0.06
Demographics							
Age, weeks					1.43	(1.25, 1.63)	<0.001
Sex (Female)					1.75	(0.99, 3.06)	0.052

* Results from logistic regression (not GEE adjusted)

Table 4

Summary of Results from Multivariable Models of Feeding

Variable in Model	Generalized Linear Model (GEE adjusted) for outcome of Feeding			Proportional Odds Model for outcome of categorized feeding gradient *		
	Adjusted Odds ratio (aOR)	95% CI for aOR	p-value	Adjusted Odds ratio (aOR)	95% CI for aOR	p-value
<i>Genes (expressed vs. unexpressed)</i>						
<i>PLXNA1</i>	2.89	(1.47, 5.67)	0.002	1.73	(0.94, 3.19)	0.08
<i>AMPK</i>	3.21	(1.09, 9.48)	0.03	1.53	(0.59, 4.02)	0.38
<i>WNT3</i>	0.59	(0.33, 1.07)	0.09	0.69	(0.42, 1.15)	0.16
<i>NPY2R</i>	0.60	(0.36, 1.002)	0.05	0.62	(0.39, 0.95)	0.03
<i>NPHP4</i>	0.59	(0.344, 1.03)	0.06	0.71	(0.43, 1.16)	0.17
<i>Demographics</i>						
PCA (in weeks)	1.43	(1.25, 1.63)	<.0001	1.37	(1.25, 1.5)	<.0001
Female Sex (vs. male)	1.75	(0.99, 3.06)	0.05	1.55	(0.99, 2.41)	0.05

* Feeding gradient: 1= 0% to 20%, 2=20% to 40%, 3=40% to 60%, 4=60% to 80%, 5=80% to 100%

Table 5

Parameter Estimates from Multivariable Models of Feeding

Parameter	Generalized Linear Model (GEE adjusted) for outcome of Feeding		Proportional Odds Model for outcome of categorized feeding gradient	
	Beta Coefficient	Standard Error of Beta	Beta Coefficient	Standard Error of Beta
<i><u>Intercept term(s)</u></i>				
Intercept	−14.41	2.64	n/a	n/a
Intercept for gradient 81 to 100	n/a	n/a	−11.69	1.76
Intercept for gradient 61 to 80	n/a	n/a	−11.36	1.75
Intercept for gradient 41 to 60	n/a	n/a	−11.09	1.74
Intercept for gradient 21 to 40	n/a	n/a	−10.65	1.73
<i><u>Genes (expressed vs. not)</u></i>				
PLXNA1	1.06	0.34	0.55	0.31
AMPK	1.17	0.55	0.43	0.49
WNT3	−0.51	0.29	−0.37	0.26
NPY2R	−0.5	0.26	−0.49	0.22
NPHP4	−0.52	0.28	−0.35	0.25
<i><u>Demographics</u></i>				
PMA (in weeks)	0.36	0.07	0.32	0.05
Female Gender (vs. male)	0.56	0.2	0.44	0.23