Supplementary for: Nutritional geometry of mitochondrial genetic effects on male fertility

M.F. Camus[1], J. Moore[1], M. Reuter[1]

[1] *Research Department of Genetics, Evolution and Environment, University College London, Gower Street, London, WC1E 6BT, United Kingdom*

Nutritional Geometry

We designed our experiments using the nutritional geometric framework. This framework aims to untangle the complex interaction between the different macromolecules in food that drive phenotypic responses. Consequently, for this study we chose to understand the relationship between protein and carbohydrate acting on mtDNA-specific male fitness. In line with the framework, we chose four defined diets ("nutritional rails") that differed in their protein-to-carbohydrate ratios (**Figure S1A**). The use pre-determined diets locks in experimental flies to each nutritional rail; if flies consume more food, they would move up the rail whereas if they consume less the datapoint would move down the rail. Consequently, the more rails are used in the study the greater the nutritional landscape is surveyed.

The second requirement for nutritional geometry is that a phenotypic trait for each individual on the nutritional landscape is collected. There is no constraint on what this phenotypic trait can be, but for our study we have chosen male fitness. Hence, for each datapoint on the nutritional landscape (**Figure S1A**) we have three phenotypic measurements: protein consumption, carbohydrate consumption and male fitness. Using non-parametric thin-plate splines, we can visualise the Z-axis (fitness) for each data point on the nutritional landscape (**Figure S1B-C**). This visualisation is in the form of a heatmap, where areas in red contain datapoints with high fitness measures and blue with low fitness. In conclusion, framework is able to dissect this complex relationship much better than a standard diet treatment as it takes into account part of the behavioural aspect, how much food an individual chose to consume to find the ideal nutritional conditions that maximise a particular phenotype.

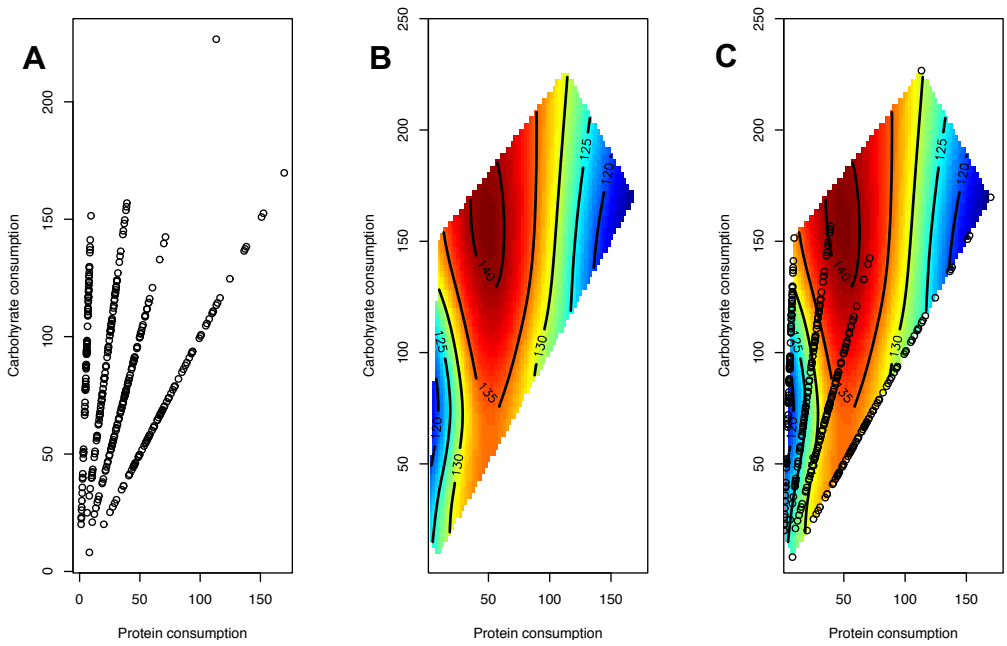

Figure S1: Principles of the nutritional geometry framework. (A) Datapoints for nutrient consumption plotted for each individual in the experiment. Given individuals are locked to experimental diets, it is possible to visualise the nutritional rails for each diet. (B) For each datapoint, we have a phenotype linked with nutritional consumption (in our case its male fertility), and we can superimpose thin-plate splines to create a fitness landscape. (C) Final figure showing nutrient consumption and phenotypic landscape.

A complete introduction to the Nutritional Geometry Framework can be found in: Simpson & Raubenheimer (1995) "The Geometric Analysis of Feeding and Nutrition: A User's Guide", *J. Insect.Phys*

Statistical Models

Diet consumption

To examine whether the genotypes varied in the quantity they consumed of each diet, we used a model building approach previously used in Camus et al. (2017). The basic model (Model C1) expressed total diet consumption (microliters) as a function of diet, with block as a random effect. To this model, we then include the random effect of mitochondrial genotype to examine if different genotypes consume on average different amount of diet (Model C2). Our final model (Model C3) adds diet as a random, genotype-specific intercept to investigate genetic variance for diet-specific consumption.

Models were compared using parametric bootstrap [1], an approach that uses simulations of data under a simpler model to generate a null distribution for the test statistic (the log-likelihood ratio) against which to compare the observed statistic.

Comparison 1 (Model C1 vs Model C2)

	LLR	df	p-value
PBtest	2.461	1	0.04811

Comparison 2 (Model C2 vs Model C3)

	LLR	df	p-value
PBtest	15.941	9	0.001015

We additionally calculated the amount of variance explained by each model; plus how much extra variance is explained by increasing the complexity of our models. This was done using the *r.squaredGLMM* function within the MuMIn package [2] in R . Model C1 explains 19.21% of the variance and is increased by 1.38% when adding extra terms (Model C2). Furthermore, this is increased again by 4.24% when adding the terms specific to Model C3.

Nutritional Requirement

We compared a base model (Model F1) that describes the experimental response surface with fixed effects for the linear, quadratic and cross-product effects of the consumed diet components, as well as a random effect for experimental block, with a more complete model (Model F2) that assess genetic variation for diet effects on fitness. For this, we added random intercept terms to Model F1 that describe how the flies of different mitochondrial genotypes vary in their average fitness (across all dietary regimes). The full model (Model F3) builds upon Model F2 an adds the genotypic-specific responses with linear, quadratic and cross-product effects of carbohydrate and protein intake. Model F3 thus allows for genotype-specific shape specification of the fitness response surface.

Comparison 1 (Model F1 vs Model F2)

	LLR	df	p-value
PBtest	9.8427	1	0.001812

Comparison 2 (Model F2 vs Model F3)

	LLR	df	p-value
PBtest	9.7451	14	0.0495

We also calculated the amount of variance explained by the fitness models. Models F1, F2 and F3 accounted for 16.62%, 20.29% and 26.4% of the variance respectively.

Further Nutritional Requirement Analysis – Reaction Norms

In addition to using the Geometric Framework to analyse our nutritional requirement data, we used linear mixed models to analyse our data. This analysis follows the same principles as the diet consumption analysis, with fitness as a response variable instead of consumption. We do have to note that it does not take into account effects due to liquid diet consumption. The basic model (Model R1) expressed total number of offspring as a function of diet, with block as a random effect. To this model, we then include the random effect of mitochondrial genotype to examine if different genotypes consume on average different amount of diet (Model R2). Our final model (Model R3) adds diet as a random, genotype-specific intercept to investigate genetic variance for diet-specific consumption.

Although we do not take diet consumption into account in this model, we still find significant mitochondrial genetic variance (Comparison 1) and mito × diet interactions (Comparison 2). This further supports our nutritional geometry results.

Comparison 1 (Model R1 vs Model R2)

	LLR	df	p-value
PBtest	3.502	1	0.001185

Comparison 2 (Model R2 vs Model R3)

	LLR	df	p-value
PBtest	3.0885	9	0.004032

We also calculated the amount of variance explained by these models. Models R1, R2 and R3 accounted for 8.55%, 11.96% and 15.72% of the variance respectively.

Permutations Tests

We wanted to determine to what degree variation in fitness across genotypes/diets was due to behavioural (variation in amount of nutrient) or physiological responses (variation in metabolic machinery). We used a permutation approach, because it will break any associations between behavioural and physiological responses to the different diets. If the variation in fitness is determined by the amount consumed or by a matching of behavioural responses with physiology, then the permutation of consumption data should lead to a lower average predicted fitness and reduced variation in fitness between genotypes. For our test, consumption values across genotypes were permutated and used to calculate precited fitness values based on the final model fitted to the data (Model F3). This was done separately for each block and dietary composition. In total, we generated 1000 datasets with permuted consumption data and comparing the distributions of means and variances in fitness across permutations to observed values of these parameters in the original data. P-values were calculated as the proportion of parameter values calculated from the permuted data that equalled or exceeded the values observed in the original dataset. Permutation tests were performed on the entire dataset.

Table S1: Essential and non-essential amino acid stock solutions.

Amino acid stock solution	**(g/200 ml)**
Essential amino acid	
F (L-phenylalanine)	**3.03**
H (L-histidine	**2.24**
K (L-lysine)	**5.74**
M (L-methionine)	**1.12**
R (L-arginine)	**4.70**
T (L-threonine)	**4.28**
V (L-valine)	**4.42**
W (L-tryptophan)	**1.45**
Non-essential amino acid	
A (L-alanine)	**5.25**
D (L-aspartate)	**2.78**
G (glycine)	**3.58**
N (L-asparagine)	**2.78**
P (L-proline)	**1.86**
Q (L-glutamine)	**6.02**
S (L-serine)	**2.51**

Table S2: Recipe for 200ml of protein solution

			Total volume 200ml
	L-ile	Powder	348mg
	L-leu	Powder	492mg
	L-tyr	Powder	252mg
	cholesterol	20mg/ml in EtOH	3ml
	CaCl2	1000x	200ul
	MgSO4	1000x	200ul
	CuSO4	1000x	200ul
	FeSO4	1000x	200ul
	MnCl2	1000x	200ul
	ZnSO4	1000x	200ul
	H_2O		Up to 50ml
	Total volume before autoclaving		50 ml
	buffer	10x acetate buffer base	20ml
	nucl/lipid soln	125x stock	1.6ml
	Yaa solutions	essential amino acid stock solution (EAA)	18.154ml
		non essential amino acid stock solution (NEAA)	18.154ml
		Na glutamate solution (100mg/ml)	5.464ml
		Cys solution (50mg/ml)	1.584ml
	Vitamin stock	47.6x stock	4.2ml
	folic acid stock	1000x stock	200ul
	Propionic acid		1.2ml
	Nipagin	100 g/l stock in 95% EtOH	3ml
	Make to total volume of 200ml with H_2O		

Table S3: Recipe for 200ml of carbohydrate solution

			Total volume 200ml
	sucrose	To match protein 1:1	6.5g
	cholesterol	20mg/ml in EtOH	3ml
	CaCl2	1000x	200ul
	MgSO4	1000x	200ul
	CuSO4	1000x	200ul
	FeSO4	1000x	200ul
	MnCl2	1000x	200ul
	ZnSO4	1000x	200ul
	H_2O		Up to 50ml
Total volume before autoclaving			50ml
	buffer	10x acetate buffer base	20ml
	nucl/lipid soln	125x stock	1.6ml
	Vitamin stock	47.6x stock	4.2ml
	folic acid stock	1000x stock	200ul
	Propionic acid		1.2ml
	Nipagin	100 g/l stock in 95% EtOH	3ml
Make to total volume of 200ml with H_2O			

References

1. Halekoh, U., and Højsgaard, S. (2014). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models: The R package pbkrtest. Journal of Statistical Software *59*, 32.
2. Burnham, K.P., and Anderson, D.R. (2003). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, (Springer New York).