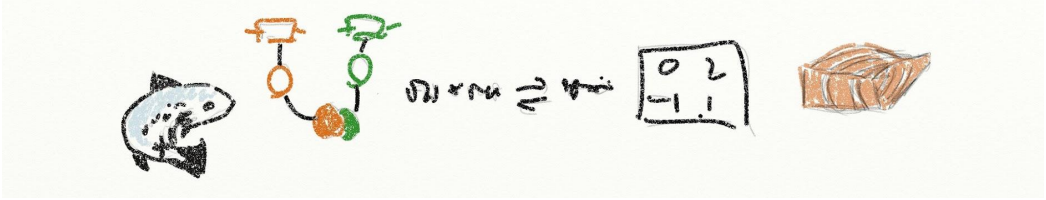
Towards the DIGITAL SALMON



Jon Olav Vik, Norwegian University of Life Sciences, Ås

Digital tools and methods in life sciences. Life Science Cluster breakfast meeting 2017-02-09

jonovik@gmail.com www.nmbu.no/prosjekter/digisal facebook.com/thedigitalsalmon

I will talk about how digital methods will transform food production.
Adopting methods that are already being applied in human medicine and industrial microbiology.

Presented at https://tlsc.no/event/digital-tools-methods-life-science/

The farming…

Photo: Shutterstock

…of salmon…

Photo: Shutterstock

…is Norway's biggest export industry besides oil, …

Photo: Shutterstock

Olafsen *et al Rep R Nor Soc Sci Lett Nor Acad Technol Sci* (2012)
http://www.sintef.no/home/Fisheries-and-Aquaculture/News/Value-created-from-productive-oceans-in-2050

…and it faces *huge* challenges.

Photo: Shutterstock

Rapid response to new challenges in salmon farming

Sustainable food production • Disease • Preserving wild salmon

Of sustainability, disease outbreaks, and preserving the wild salmon.
 Today I will outline how we are addressing this first one.

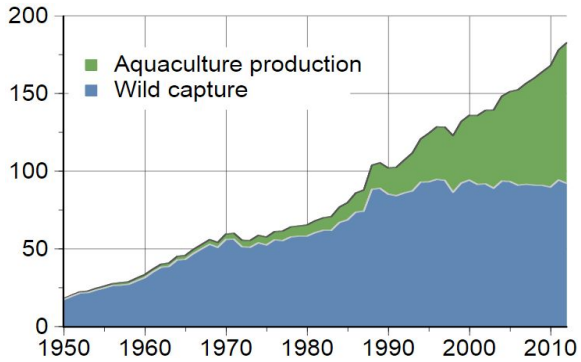
Salmon are carnivores by nature, and in the early years of salmon farming, they were fed something…*similar* to their natural diet…

Photo: Shutterstock (salmon, sand eels), Wikipedia (herring, prawn, Gammarus)

…until about 15 years ago.

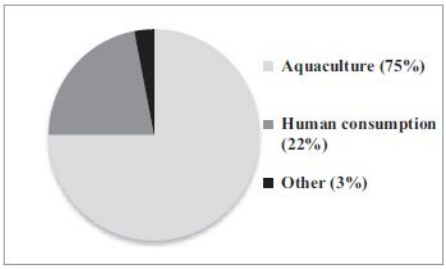
Photo: Shutterstock

Fish oil is now a scarce resource



Global harvest of aquatic organisms in million tonnes, 1950–2010, as reported by the FAO

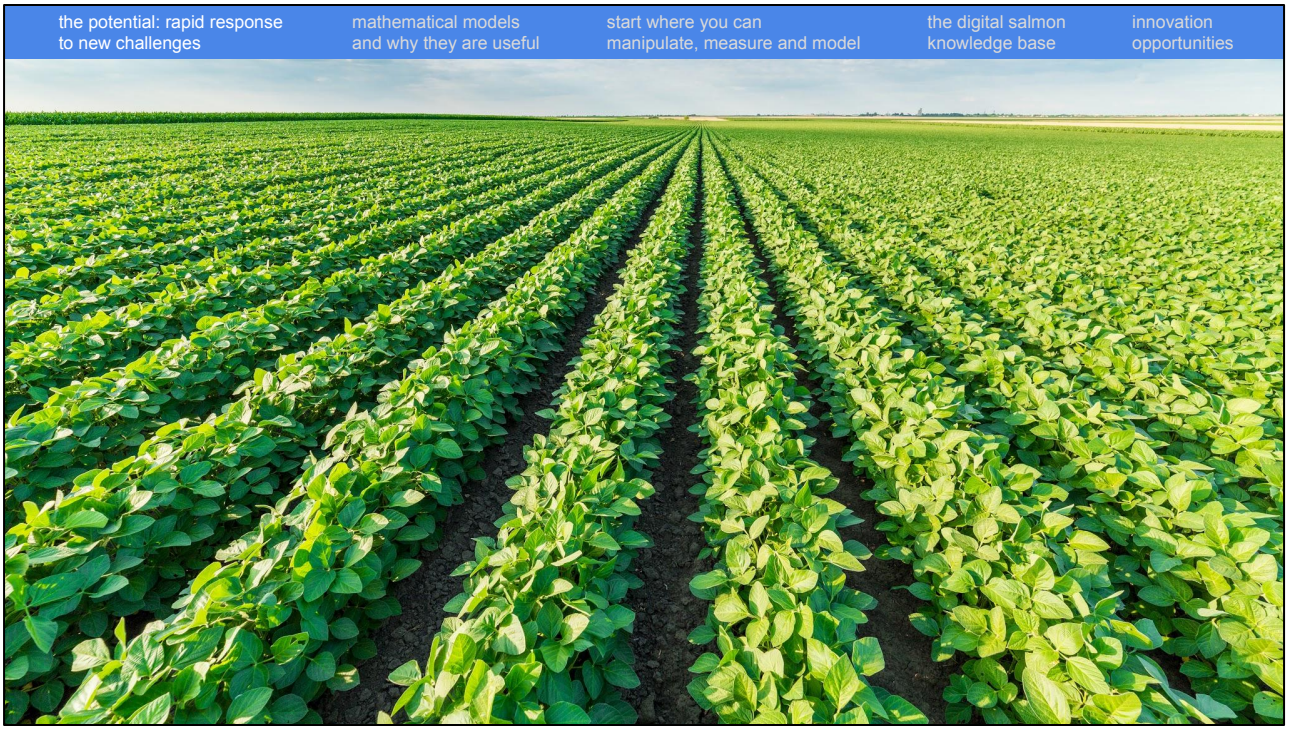
Fish oil usage in 2012



D.R. Tocher / Aquaculture 449 (2015) 94–107

9

But now fish oil is a scarce resource. So instead, we taught salmon to eat…



…*this*. Today, about 75% of the fat and protein…

Photo: Shutterstock

…comes from plants. But it's not straightforward. Over a period of ten years, salmon feed has changed from containing ten ingredients with one protein source and one fat source, to more than thirty ingredients with several protein and fat sources.

Feedstuff prices fluctuate rapidly, and it is time-consuming and expensive to trial new diets on fish.

Also, plant-fed salmon is not particularly sustainable, because salmon now competes with plant production for human food.
Therefore, researchers are now exploring novel, sustainable feed ingredients such as yeast, bacterial meal and microalgae.

Photo: Shutterstock

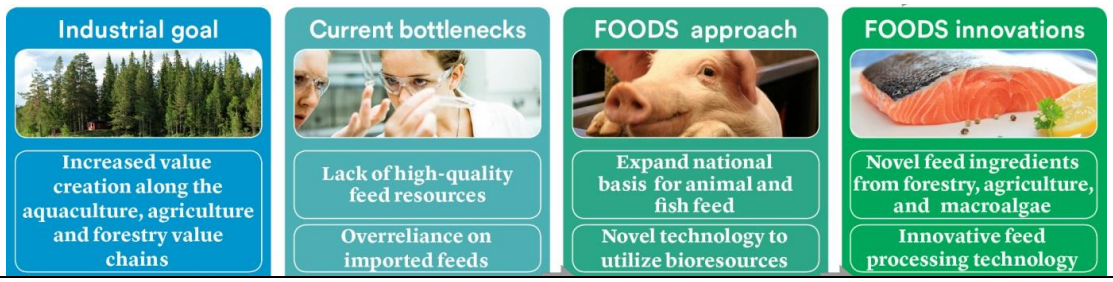
Eat this: Foods of Norway

Margareth Øverland

Centre for Research-based Innovation at NMBU (2015-2023)

Sustainable feed ingredients from natural bioresources that are not suitable for direct human consumption
e.g. yeast, bacterial meal, microalgae

http://www.foodsofnorway.net



The newly started Centre for Research-driven Innovation, called *Foods of Norway*, is evaluating these novel feedstuffs in salmon -- and pigs and chickens.

But still it is hard!

Many alternative feed ingredients

New options arise frequently

Rapidly shifting prices

Genetic variation in feed utilization

Optimization wishlist: high-quality meat, low cost, sustainable feeds

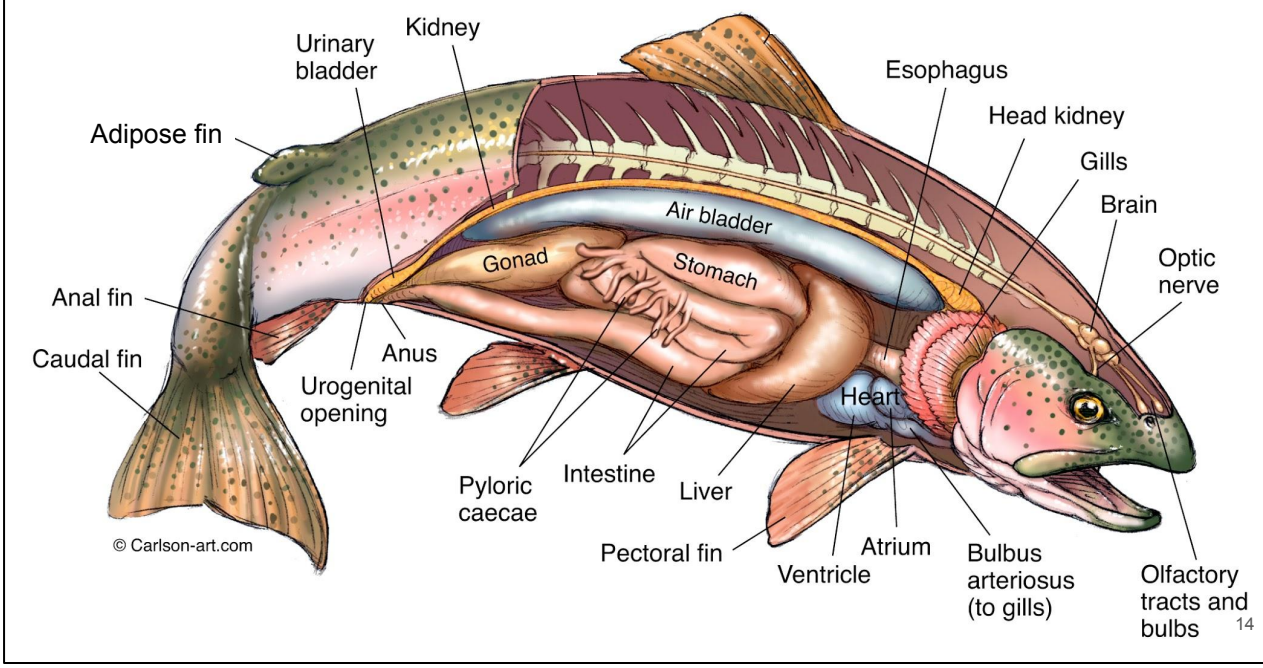
Feeding trials are slow and expensive

So there are many promising, novel, sustainable feedstuffs becoming available.

But composing a good recipe from the available ingredients remains a hard problem.

All the more so because this is not salmon's natural diet.

And the body's responses to a diet are complex and involve many organs.



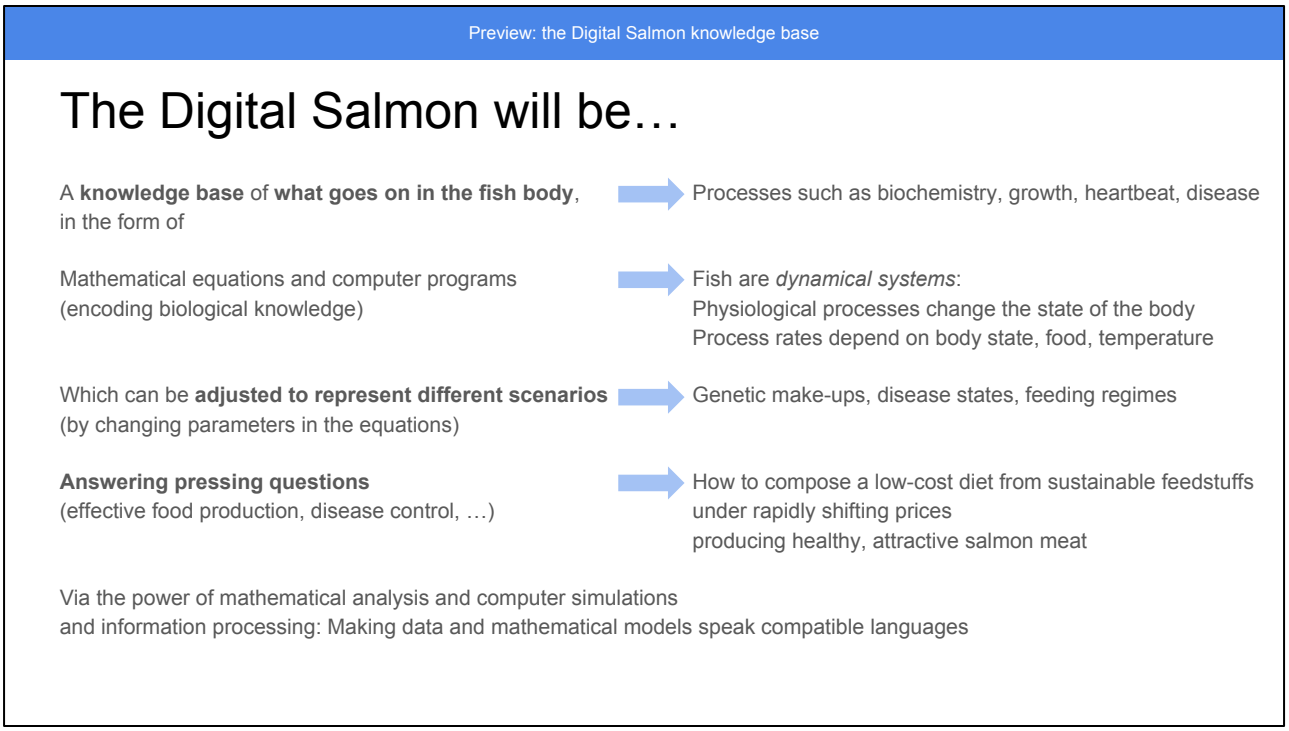
And this brings us to systems biology, understanding the living body as a set of components that both affect each other and depend on each other.

This allows much more powerful deductions from experiments than simply viewing the input-output relationship as a black box.

We seek to account for the *mechanisms* that underlie the relationship between feed input and fillet output.

And the way to do this accounting is via mathematical models.

http://www.biographixmedia.com/biology/trout-fish-anatomy.html



The *name* "Digital Salmon" is a vision that has just started to become reality.

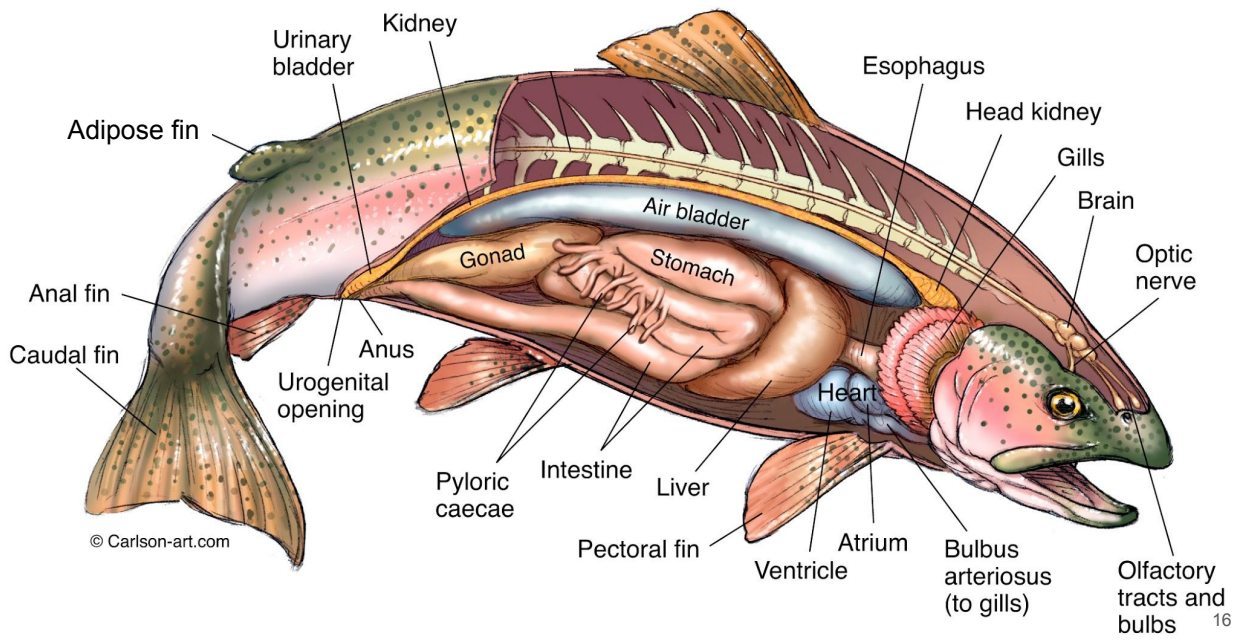
The idea is to build a tightly integrated knowledge base of salmon physiology,

in the form of *mathematical models* linked to omics data:

genomics, gene expression, metabolomics and so on.

It will help the salmon farming industry to navigate conflicting demands -- of sustainability, shifting feed prices, diseases and product quality. The industry needs to develop a flexible, integrated basis of knowledge for rapid response to new challenges.

I'll try to illustrate what this means in practice.

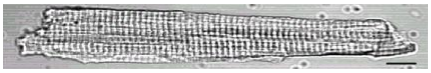


Let me use the heartbeat as a first example of mathematical modelling, even though it's not particularly related to feeding…

http://www.biographixmedia.com/biology/trout-fish-anatomy.html

The unique benefits of mathematical modelling

A *purposeful* simplification of reality
Highlights similarities (form of equations)
and differences (values of parameters)
"Parameters are phenotypes"
Connects knowledge into a functional whole



Heart muscle cell
0.13 mm long
(Louch et al 2004)

Differential equations
Process rate is a function of state

$$dV/dt = C (I_1 + I_2 + ...)$$
$$I = g \times O \times (V - E)$$

Rate of change in voltage = (Conductance of cell membrane) × (Current 1 + Current 2 + ...)
Current = (Amount of ion channel) × (Proportion of channels that are open) × (Voltage – Equilibrium voltage)

This slide shows several models that have proved useful in understanding heart disease.

A **model** is a *purposeful simplification* of reality, designed to imitate certain phenomena or characteristics of a system while downplaying non-essential aspects. Its value lies in the ability to generalise insights from the model to a broader class of systems. Thus, a lab mouse can be a model representing mammals in general; an *in vitro* heart cell can represent the cells in an intact heart; and a set of differential equations can approximate the dynamic behaviour of a biological system.

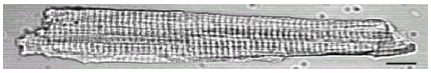
There are some unique advantages to having a *mathematical* formulation of our knowledge and hypotheses. The most obvious reason is that you can use the rules of mathematics and computer simulations to reveal previously unrecognized implications of existing knowledge or assumptions, thereby deducing new insights or disproving false hypotheses. Furthermore, having a mathematical model for a class of systems makes clear both what these systems have in common (the mathematical form of the equations) and how they differ (in their respective parameter values). The parameters of a mathematical model can serve as phenotypes that condense vast numerical phenotypic data into fewer, more informative measures.

But there are deeper benefits of mathematical modelling…

http://commons.wikimedia.org/wiki/File:Lab_mouse_mg_3140.jpg
Louch WE, Bito V, Heinzel FR, Macianskiene R, Vanhaecke J, Flameng W, Mubagwa K & Sipido KR (2004). Reduced synchrony of Ca2+ release with loss of T-tubules—a comparison to Ca2+ release in human failing cardiomyocytes. *Cardiovasc Res* **62,** 63–73.
10.1016/j.cardiores.2003.12.031 Scale bar = 10 um

The unique benefits of mathematical modelling

A *purposeful* simplification of reality
Highlights similarities (form of equations)
and differences (values of parameters)
"Parameters are phenotypes"
Connects knowledge into a functional whole



Heart muscle cell
0.13 mm long
(Louch et al 2004)

Differential equations
Process rate is a function of state

$$dV/dt = C (I_1 + I_2 + ...)$$
$$I = g \times O \times (V - E)$$

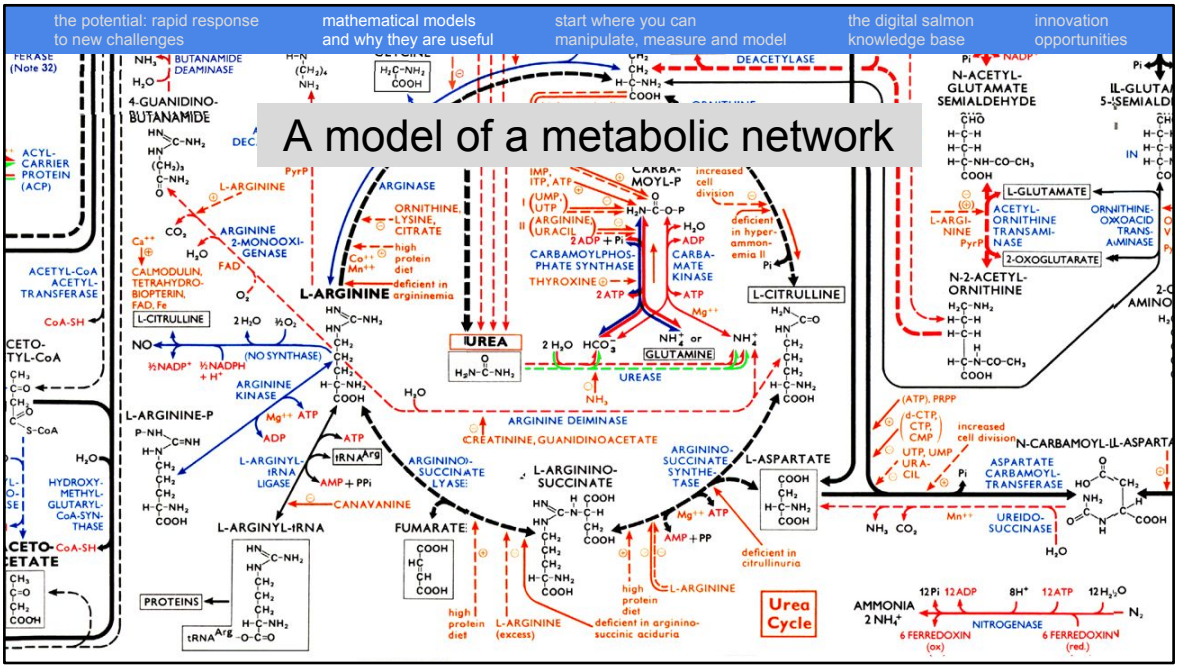
Rate of change in voltage = (Conductance of cell membrane) × (Current 1 + Current 2 + …)
Current = (Amount of ion channel) × (Proportion of channels that are open) × (Voltage – Equilibrium voltage)

But there are deeper benefits of mathematical modelling. Phrasing physiological knowledge in mathematical language enables you to connect big data into a functional whole. It forces you to be explicit about your various hypotheses. It can suggest and guide experimental and empirical work by pointing out key questions and the type of data needed. And finally, models help synthesize intellectual capital from different scientific disciplines.

The heartbeat is a prime example of a phenomenon that has been modelled on scales from molecules to the entire organ.

But the physiological process I am currently working to model, is *metabolism*.

http://commons.wikimedia.org/wiki/File:Lab_mouse_mg_3140.jpg
Louch WE, Bito V, Heinzel FR, Macianskiene R, Vanhaecke J, Flameng W, Mubagwa K & Sipido KR (2004). Reduced synchrony of Ca2+ release with loss of T-tubules—a comparison to Ca2+ release in human failing cardiomyocytes. *Cardiovasc Res* **62,** 63–73.
10.1016/j.cardiores.2003.12.031 Scale bar = 10 um



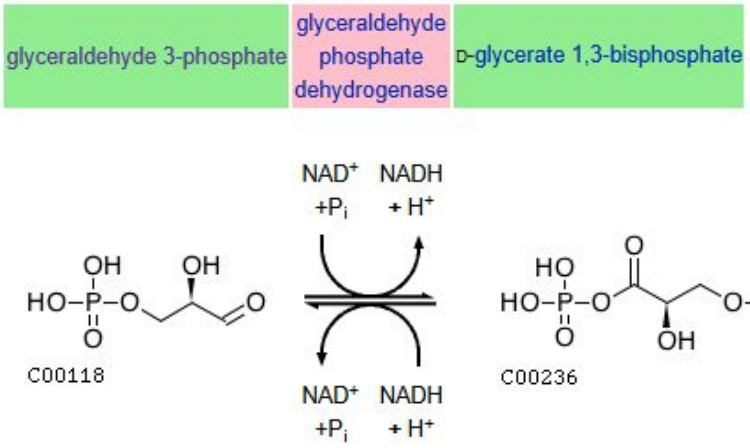
Metabolism means the biochemical reactions that go on inside you, building up and breaking down molecules to serve the body's needs.

Everything you are is built from stuff you ate.

Most of these reactions are catalyzed by very specific enzymes, produced from your genetic recipe book.
Thus, metabolism is a process that is reflected very directly in genomics and other omics data.

http://biochemical-pathways.com/

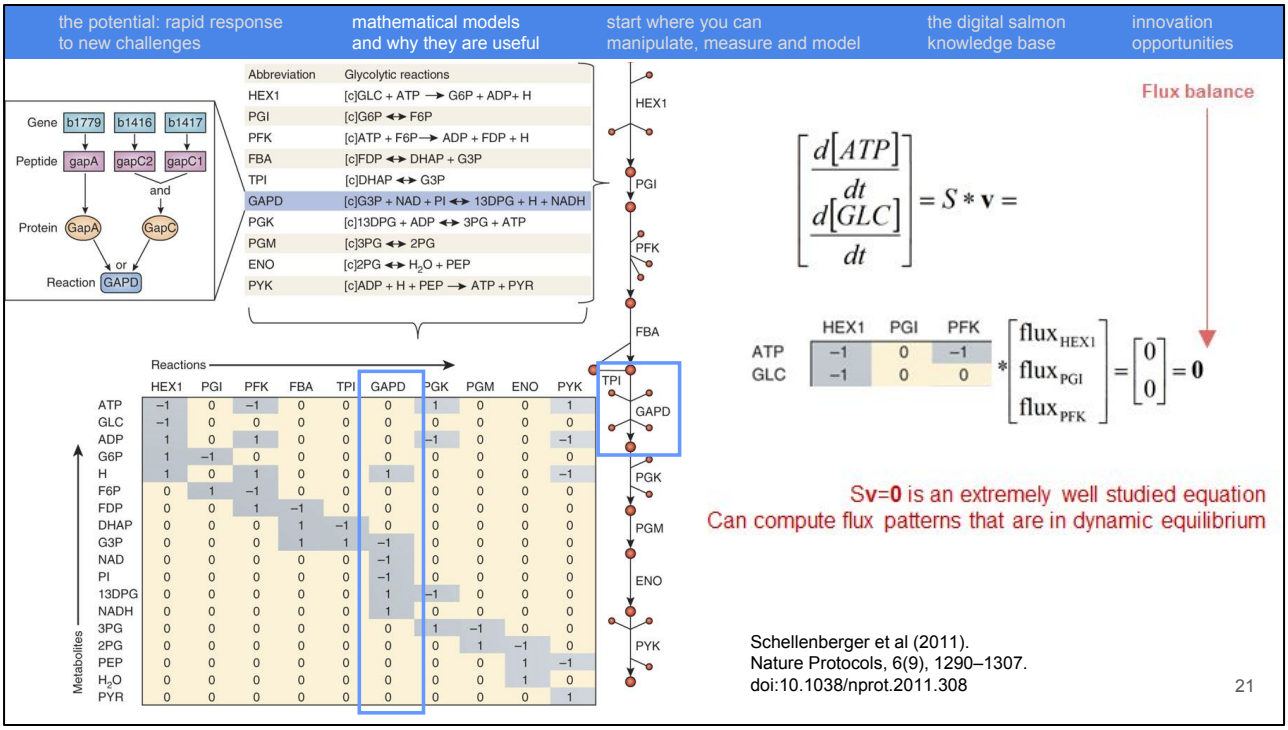
Biochemical reaction: GAPD



Compound C00118 at KEGG Pathway Database. Enzyme 1.2.1.12 at KEGG Pathway Database. Reaction R01063 at KEGG Pathway Database. Compound C00236 at KEGG Pathway Database.

Here is a reaction from the breakdown of sugar, showing how a molecule gets changed.

Modified from
https://en.wikipedia.org/wiki/Glyceraldehyde_3-phosphate_dehydrogenase
and diagram at reaction R01063 at KEGG Pathway Database.



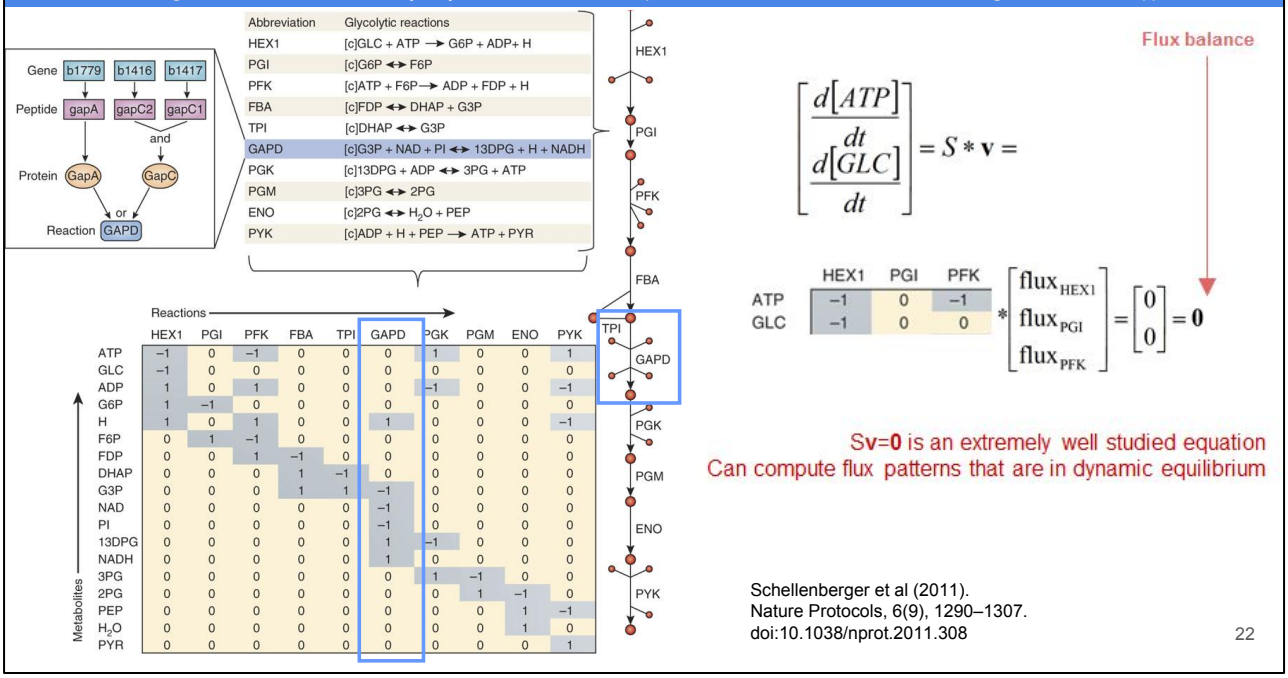
Here is the same reaction, together with the other reactions in the pathway that breaks down sugar.

You can see the same reaction in this graph…

…and in this matrix. One molecule of G3P gets consumed, along with one of NAD and another of PI. This creates one molecule of one-three-DPG and one of NADH. All the other reactions get their own column, and their own part of the graph.

Now, the *enzymes* that *catalyze* these reactions are coded for in your genome. And the *mapping* from gene to protein to reaction is the essence of metabolic modelling.

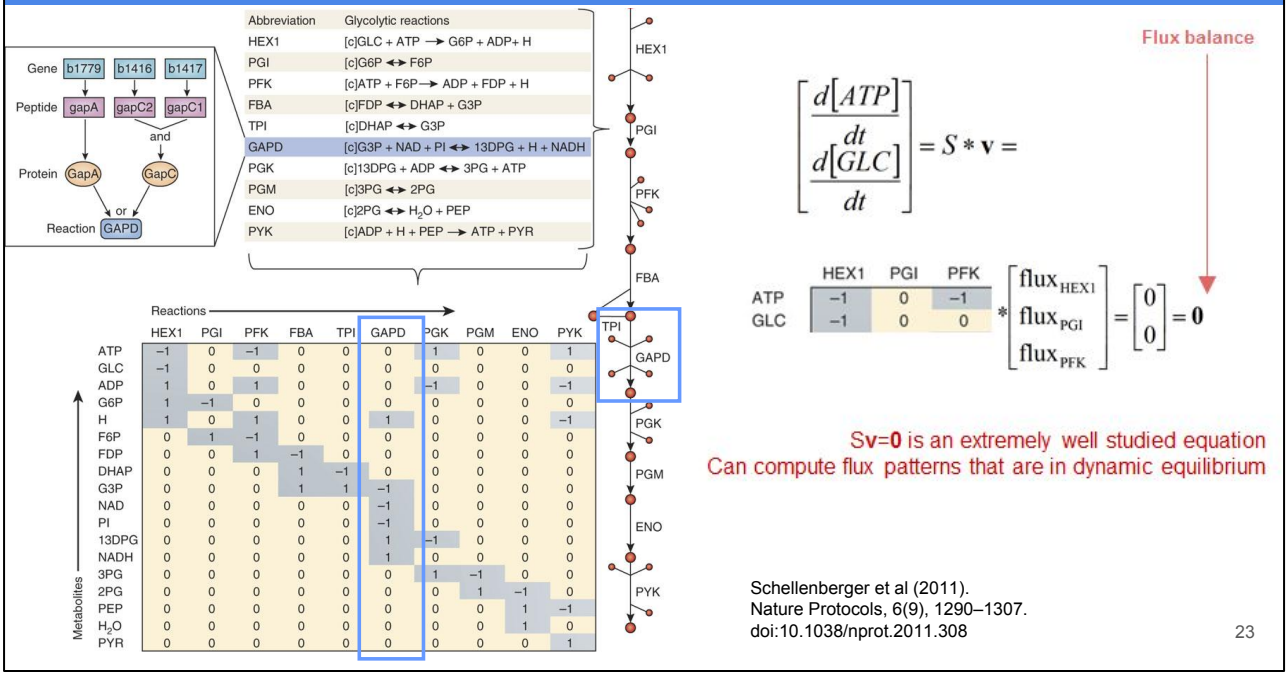
Once you have this biological knowledge in mathematical form, you can compute upon it…



Once you have this biological knowledge in mathematical form, you can compute upon it. Consider this small part of the matrix.
The rate of change of ATP concentration…is the sum of all the reaction rates that have ATP in them.
Same for glucose (GLC). You can compute all of these by multiplying this matrix with the vector of flux rates.

Now, if you assume *flux balance*, so that concentrations stay the same, you get the equation S * **v** = **0**.
This is *the* most studied equation in linear algebra. With this you can compute *all* the flux patterns that are in dynamic equilibrium.

Measurements of gene expression and metabolite concentrations enable us to…



Measurements of gene expression and metabolite concentrations enable us to mathematically model the biochemical reaction network. Gene expression tells us which reactions are active in a given tissue on a given diet, and metabolomics tells us how much we have of the various metabolites. We can then compute what are the possible flux patterns through the reaction network that are consistent with our observations. A little data will give rough constraints on what is happening, more data will give a clearer picture.

The analysis of these models will suggest hypotheses and designs for new experiments.

Integrating omics data in metabolic models

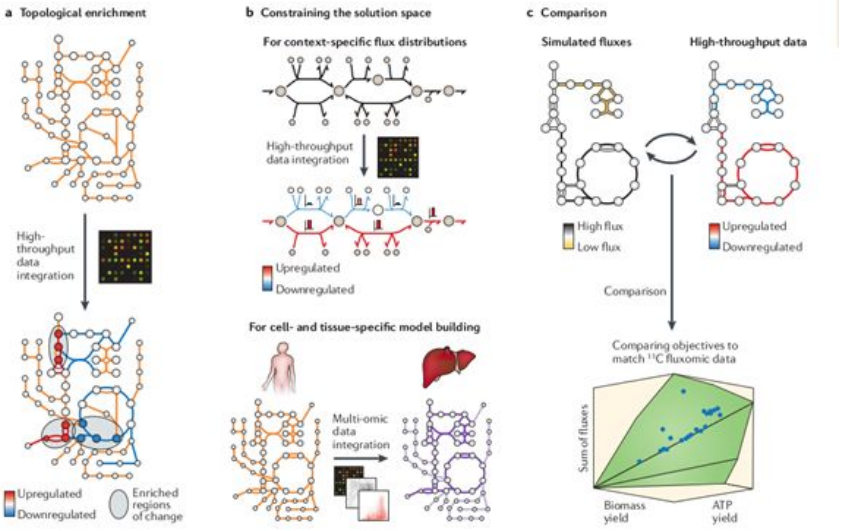
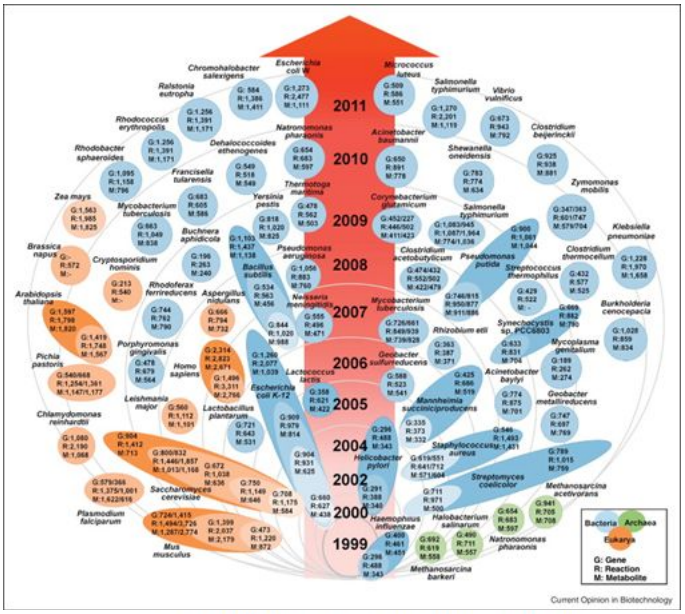


Figure 1 | **The multiple uses of high-throughput data in constraint-based models.** Constraint-based modelling can be used to interpret and augment omic data sets by using an underlying cellular network that has been biochemically validated. Metabolites are represented by circles. **a** | Similarly to pathway enrichment analysis and interaction

Bordbar et al (2014)
Nature Reviews Genetics
15:107–20
doi:10.1038/nrg3643



Kim et al (2012) Curr Opin Biotechnol 23:617-623

[Shows that constraint-based modelling of metabolic reaction networks is a tried and tested method.]

And now we are at the second question…

[And now we are at the second question…]

How to get started?

That is a question that kept me awake for a few years, until the answer presented itself.

If you have the luxury of *choosing* your biological problem,

please let it be one…

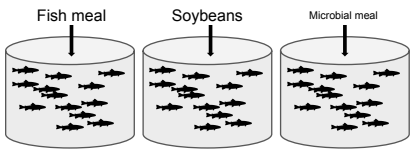
Start where you can *manipulate, measure and model*

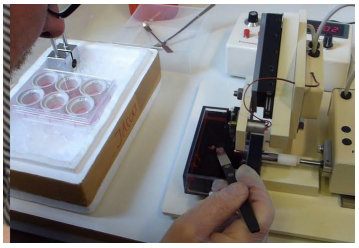
[*please* let it be one…]

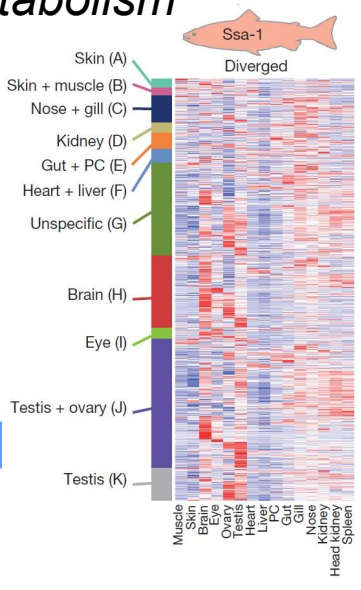
that you can manipulate, measure and model.

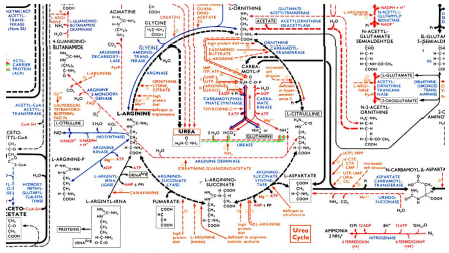
Because if *any one of these* is difficult, it's going to be hard to make sense of what is happening.

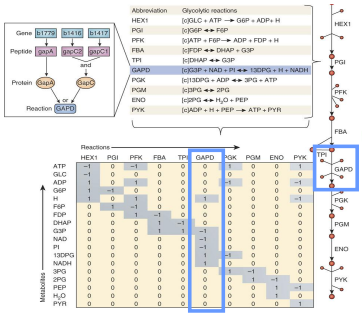
Manipulate, measure and model: *Metabolism*











The topic of our first project (http://tinyurl.com/genosysfat) is the synthesis of healthy omega-3 fatty acids.

We are currently in the process of building such a genome-scale metabolic model for salmon.

Then, we will look at tissue-specific subsets of the network, for example in liver and muscle.

https://pixabay.com/en/fish-black-fishing-silhouette-161320/

The Digital Salmon will be…

A **knowledge base** of **what goes on in the fish body**, in the form of → Processes such as biochemistry, growth, heartbeat, disease

Mathematical equations and computer programs (encoding biological knowledge) → Fish are *dynamical systems*:
Physiological processes change the state of the body
Process rates depend on body state, food, temperature

Which can be **adjusted to represent different scenarios** (by changing parameters in the equations) → Genetic make-ups, disease states, feeding regimes

Answering pressing questions (effective food production, disease control, …) → How to compose a low-cost diet from sustainable feedstuffs under rapidly shifting prices
producing healthy, attractive salmon meat

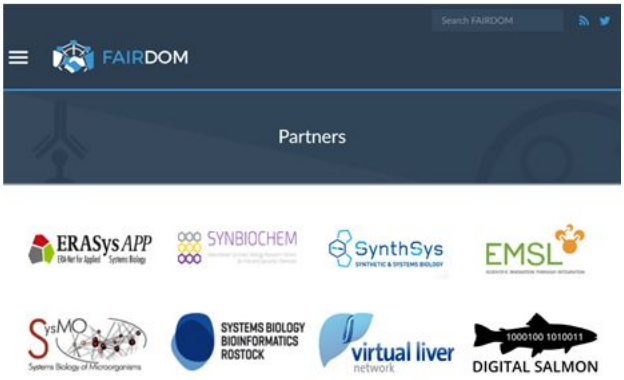
Via the power of mathematical analysis and computer simulations and information processing: **Making data and mathematical models speak compatible languages**

The Digital Salmon will help the salmon farming industry to navigate conflicting demands -- of sustainability, shifting feed prices, diseases and product quality. The industry needs to develop a flexible, integrated basis of knowledge for rapid response to new challenges.

And we will begin with the challenges of novel, sustainable feedstuffs.

Speaking compatible languages: FAIRDOM

http://fair-dom.org



The Digital Salmon is a partner in the multinational FAIRDOM initiative for common standards in bioinformatics and systems biology.

It stands for Findable, Accessible, Interoperable and Reusable

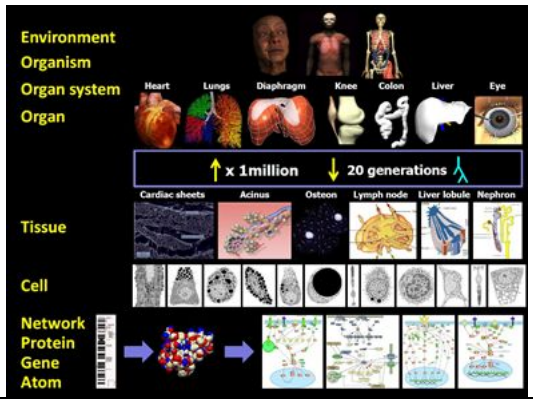
Data, Operating procedures and Models.

The road ahead: More physiological phenomena

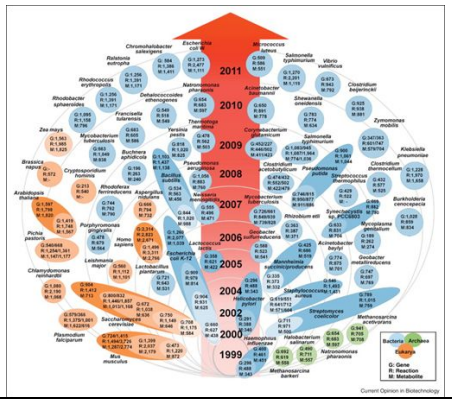
Protein uptake and growth, adapting to seawater, parasite resistance, …

Borrowing from human biomedicine (personalized medicine), industrial microbiology (biofuels, cell factories for medicines), and model organisms.

Ambitions of the
Virtual Physiological Human
Figure: Peter J. Hunter



Kim et al (2012) Curr Opin Biotechnol 23:617-623



Innovation opportunities

Model-based recommendations for feeding and breeding
– based on user data (genetics, temperature, ...)

Sustainable food production: Rapid in silico screening
of new feed ingredients for feasibility and function,
reducing the number of failed trials

Technology for "phenomics",
i.e. massive measurement of physiological characteristics

…

Consortium

Host: Norwegian University of Life Sciences — systems biology, genomics, salmon gut microbiota

Partners:

1.	Norwegian University of Science and Technology	metabolomics, systems biology
2.	University of Bergen	regulatory annotation of salmon genome
3.	University of Tromsø	metagenomics
4.	Institute of marine research	gene editing
5.	AquaGen (salmon breeding company)	salmon families, genotypic & phenotypic data
6.	Wageningen University and Research Centre, NL	metabolic reaction network modeling
7.	University of Stirling, UK	genetics of fat metabolism in salmon
8.	EWOS (feed producer)	feed composition expertise

Charles Press (NMBU/BasAM)	Torgeir Hvidsten (NMBU/IKBM)	Jacob Torgersen (AquaGen)
Arne Gjuvsland (NMBU/IHA)	Phil Pope (NMBU/IKBM)	Inge Jonassen (UiB/Informatics)
Fabian Grammes (NMBU/IHA)	Knut Rudi (NMBU/IKBM)	Eivind Valen (UiB/Informatics)
Matthew Kent (NMBU/IHA)	Lars Snipen (NMBU/IKBM)	Rolf Edvardsen (IMR)
Sigbjørn Lien (NMBU/IHA)	Thrond Haugen (NMBU/INA)	Anna Wargelius (IMR)
Liv Torunn Mydland (NMBU/IHA)	Henning Sørum (NMBU/MatInf)	Nils Peder Willassen (UiT/Chemistry)
Felipe Reveco (NMBU/IHA)	Per Winge (NTNU/Biology)	Michael Leaver (Stirling)
Simen Sandve (NMBU/IHA)	Trygve Brautaset (NTNU/Biotech)	Vítor dos Santos (Wageningen)
Jon Olav Vik (NMBU/IHA)	Per Bruheim (NTNU/Biotech)	Peter Schaap (Wageningen)
Dag Inge Våge (NMBU/IHA)	Stig Omholt (NTNU/MedFak)	Dominic Nanton (EWOS)
Margareth Øverland (NMBU/IHA)	Nina Santi (AquaGen)	

CENTRE FOR DIGITAL LIFE NORWAY

www.nmbu.no/prosjekter/digisal
facebook.com/thedigitalsalmon

Many people and institutions are involved in this project, and we welcome input from other interested parties. Visit our web page and have a look! Thank you.

Photo: Shutterstock
Lightbox (picture collection) at Shutterstock:
http://www.shutterstock.com/public_lightbox.mhtml?lightbox_id=43925734&code=a274444ce9392ea48afbb2e2b2e6c5c6