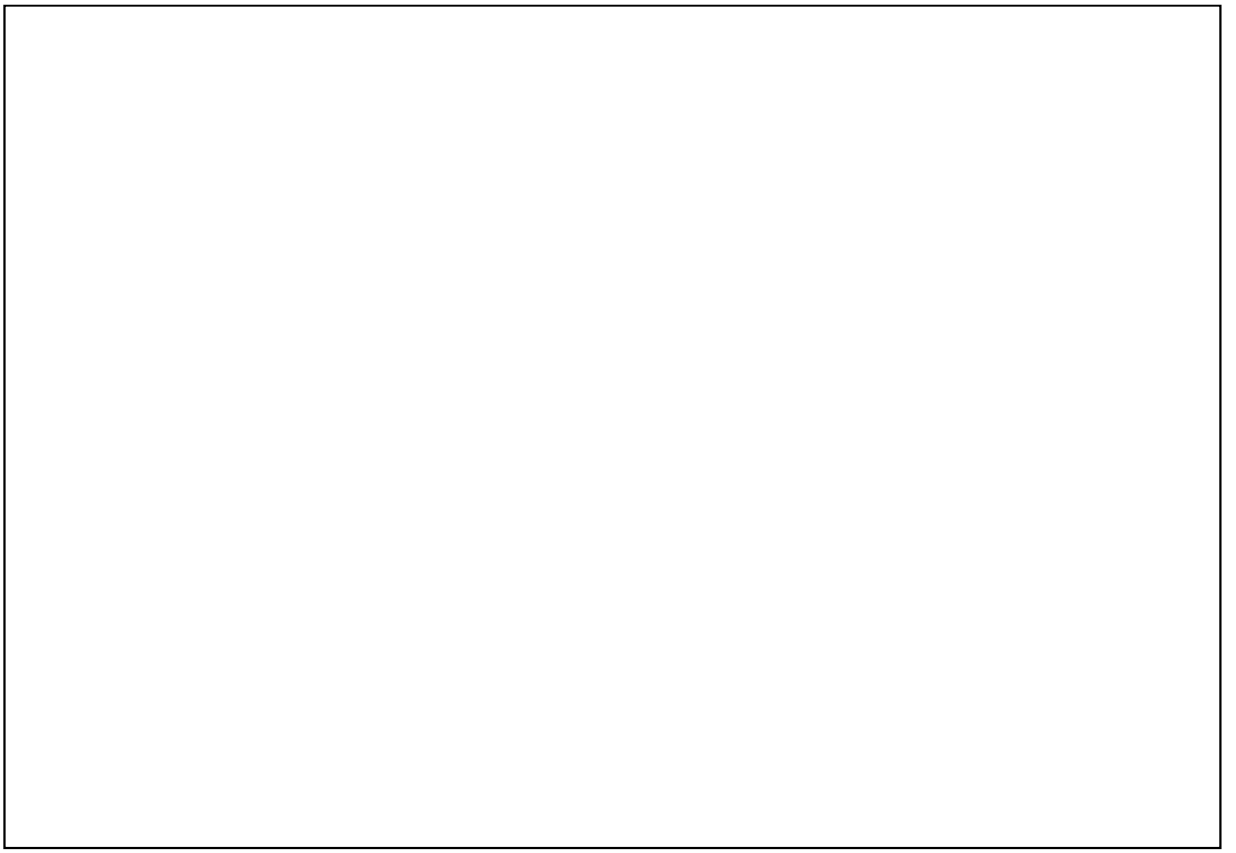


Causally-cohesive genotype-phenotype (cGP) models

First steps toward a GP map
of the single cardiomyocyte (heart muscle cell)

Jon Olav Vik, Arne B. Gjuvsland,
Stig Omholt, et al.

”Bioinformatics for molecular biology”
16.09.2009

The slide notes vary a lot in quantity and quality, and are sometimes taken slightly o

Also, I apologize for the inconsistent formatting and occasional mysteriously missin

Characteristics of a real quantitative genetics theory

- Capable of linking genes, phenotypes and population-level genetic phenomena through a causal understanding of the GP map

- Mathematical models describing how phenotypes arise from lower level processes
- Models have an articulated relation to genotypes
- Populations of dynamic systems connected to genetic maps and sequence information

Genetic variation - some problems

- Dominance (2 690 000)
- Epistasis (750 000)
- Penetrance (477 000)
- G x P interaction (386 000)
- Expressivity (318 000)
- Heterosis (286 000)
- Phenotypic plasticity (184 000)
- Pleiotropy (157 000)
- Canalization (31 800)

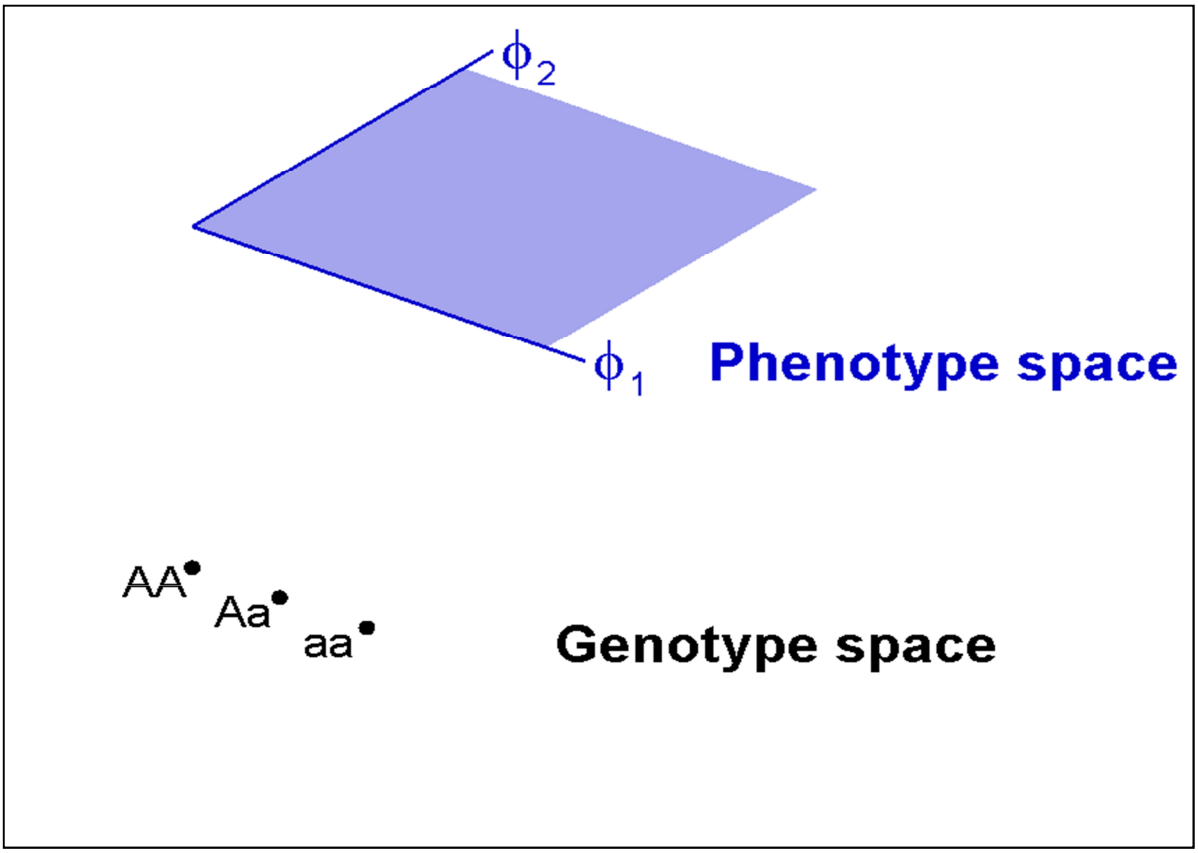
- Modularity
- Evolvability
- Developmental constraints
- Developmental dissociability
- Biological versatility
- Morphological integration

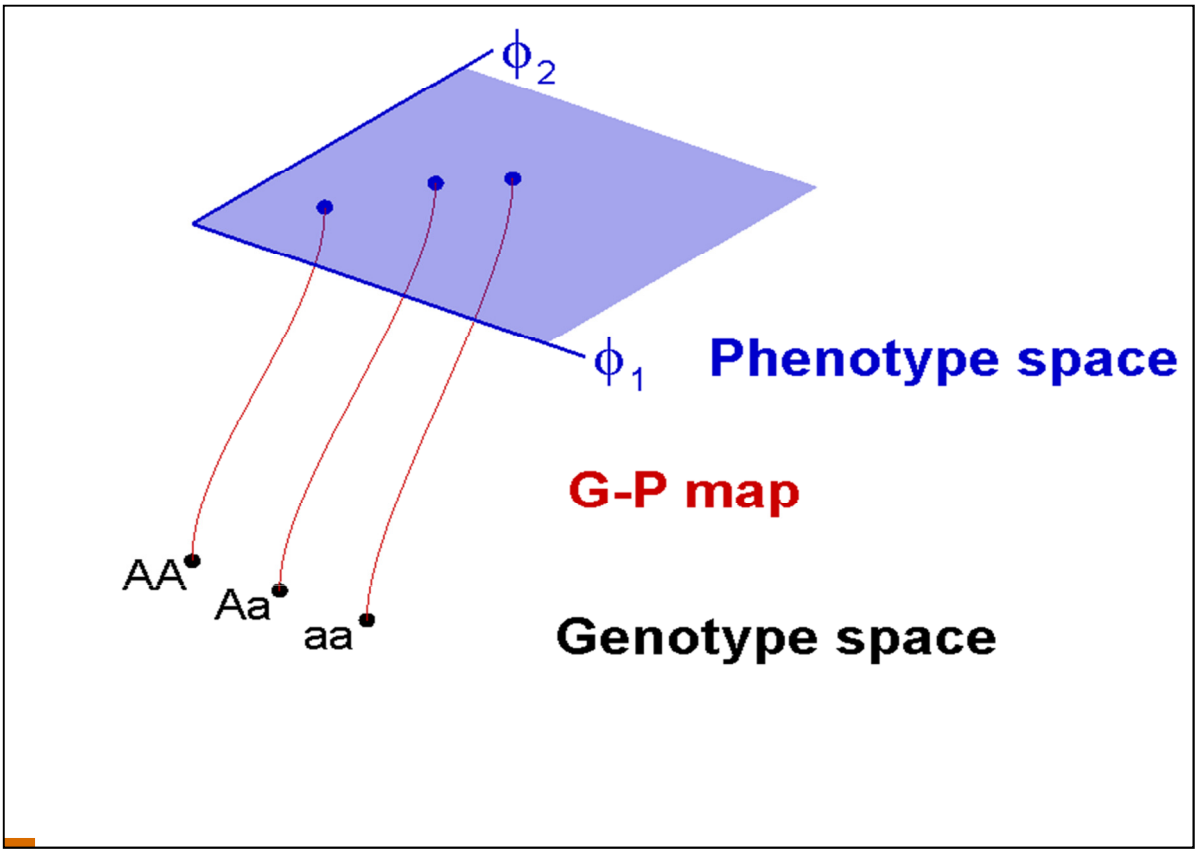
Left column: GP map phenomena we can deal with now in some sense

Right column: GP map phenomena that are still very challenging to deal with in mechanistic terms

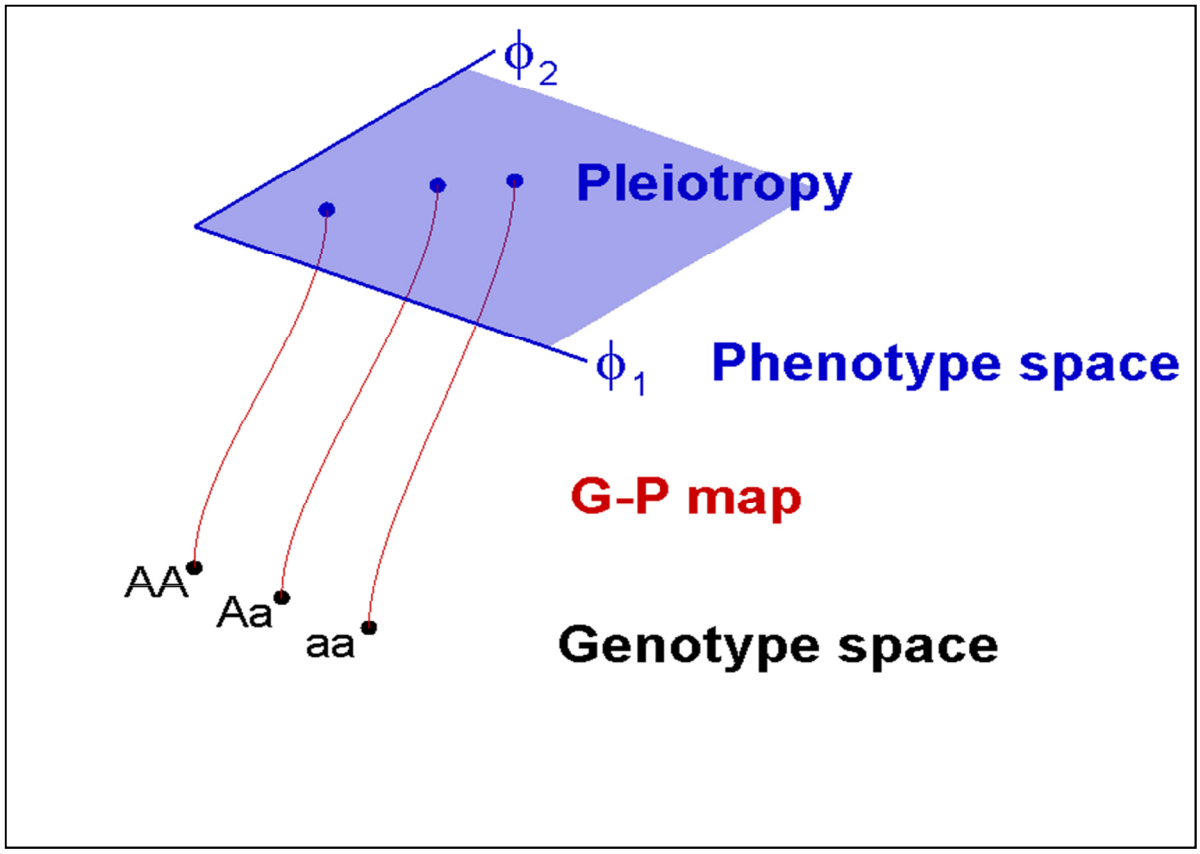
(Numbers are Google hits...)

The GP map as a conceptual meeting place

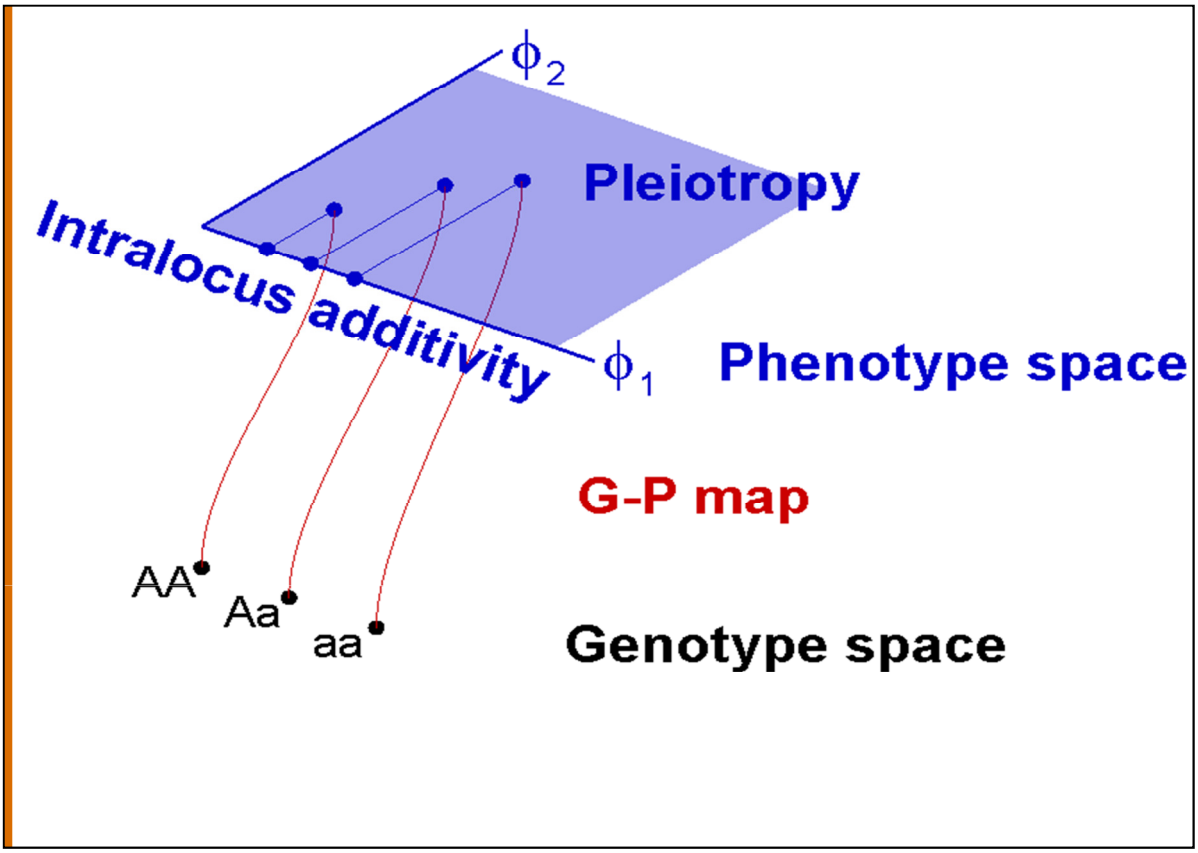

ϕ_2
ϕ_1
Phenotype space
AA
Aa
aa
Genotype space



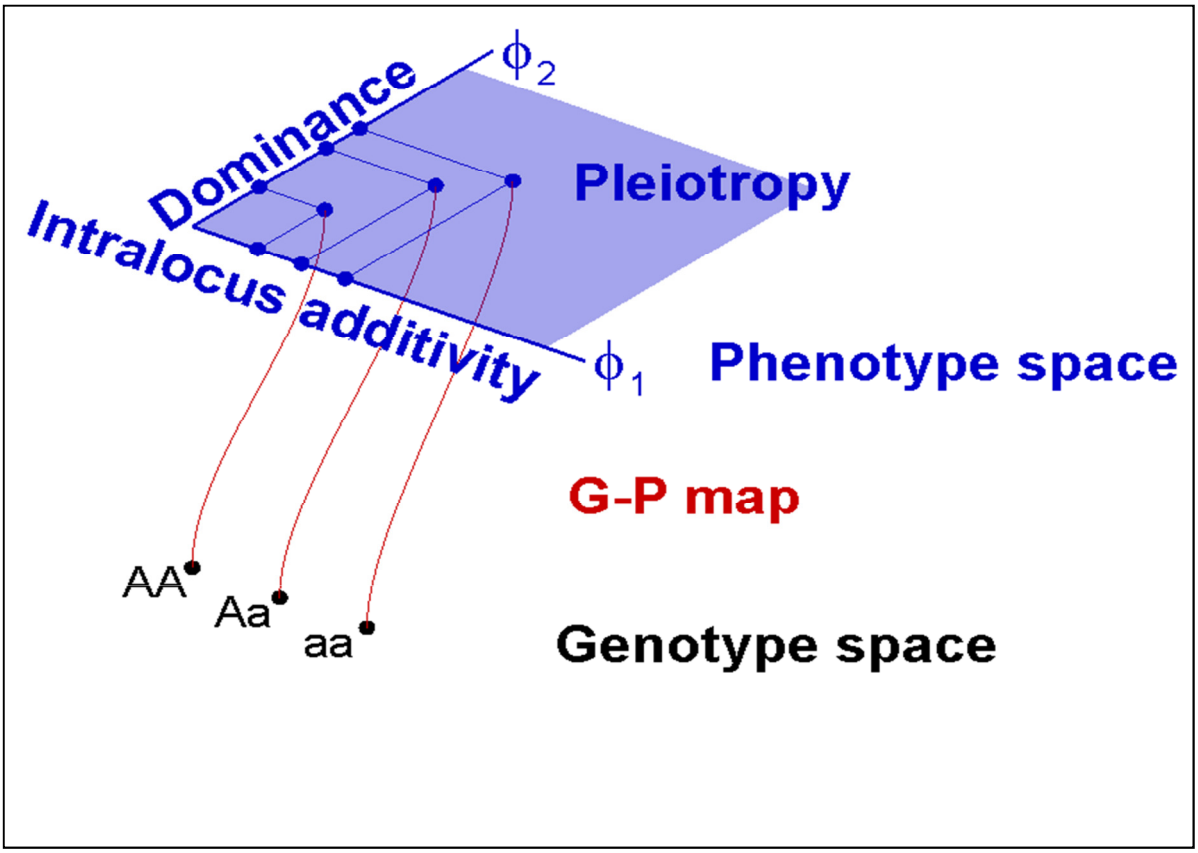
...a mapping from genotype space to phenotype space.



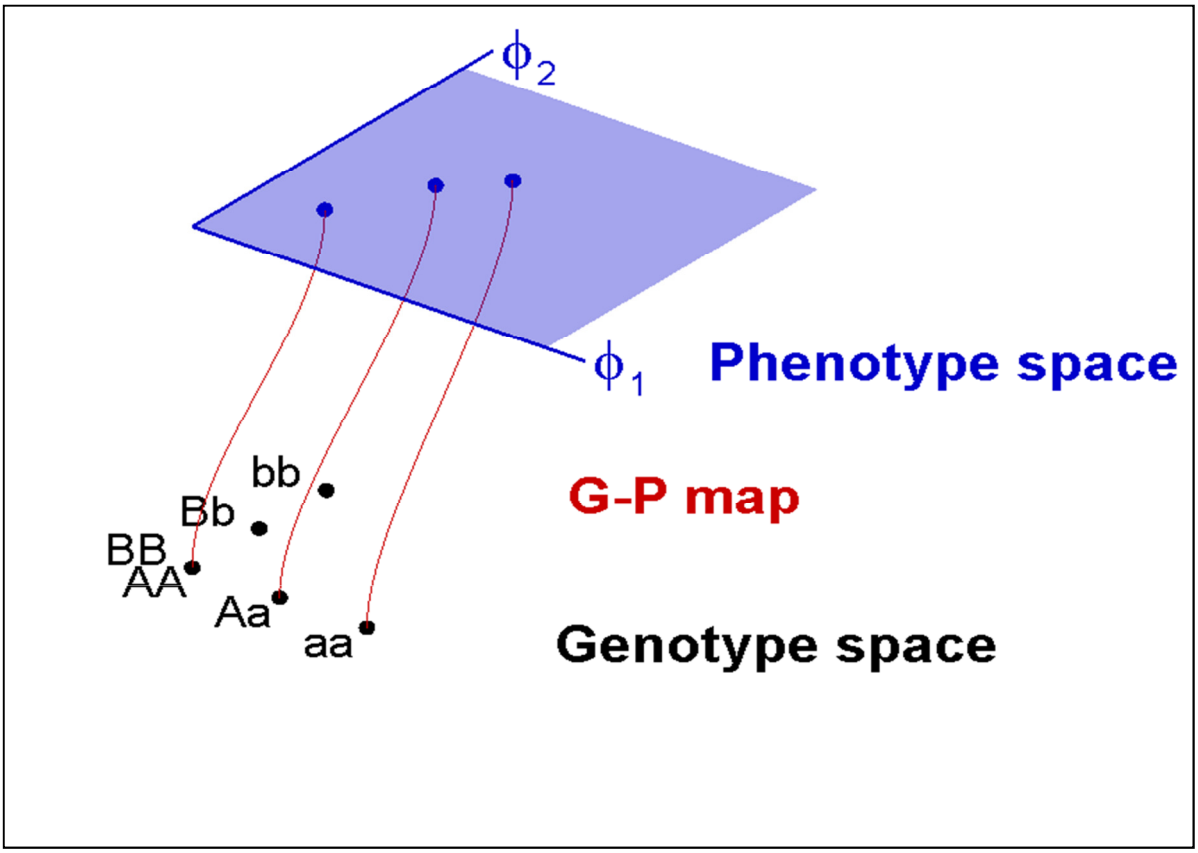
Often, a single gene affects multiple traits, a phenomenon called *pleiotropy*. [point along axes phi1 and phi2].



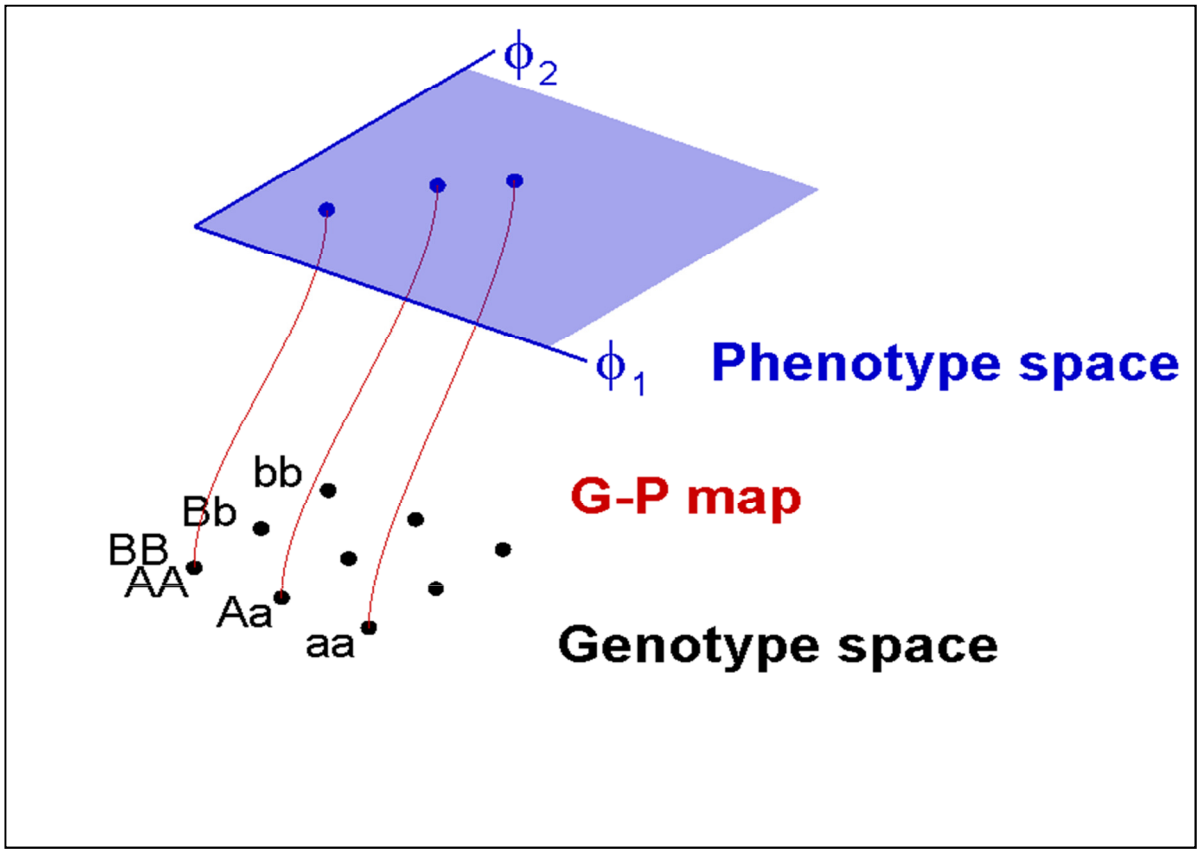
In this illustration, the action of gene *a* with respect to phenotype phi1 is additive...

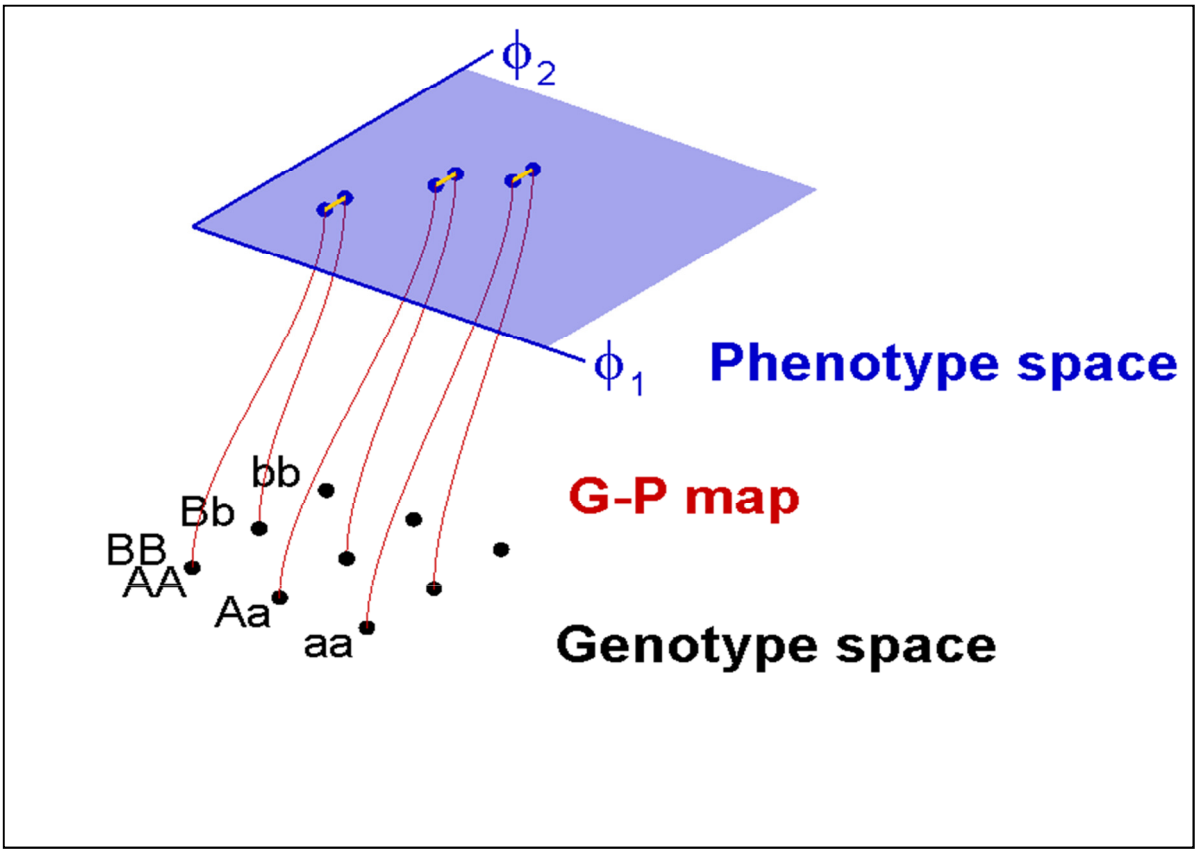


...while its effect on phenotype phi2 is nonadditive, the heterozygote being closer to the phenotype of (little a, little a).

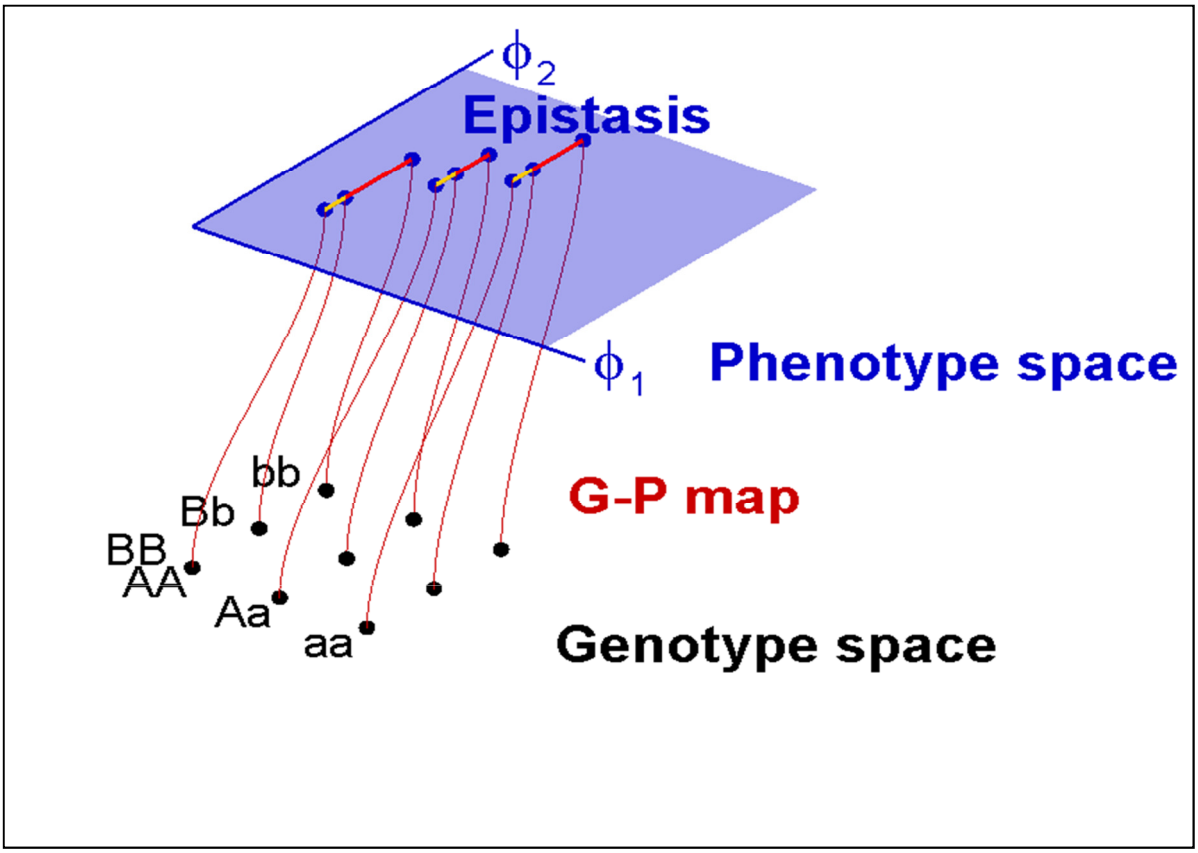


Many phenotypes are affected by several genes...


ϕ_2
ϕ_1
Phenotype space
G-P map
Genotype space
bb
Bb
BB
AA
Aa
aa

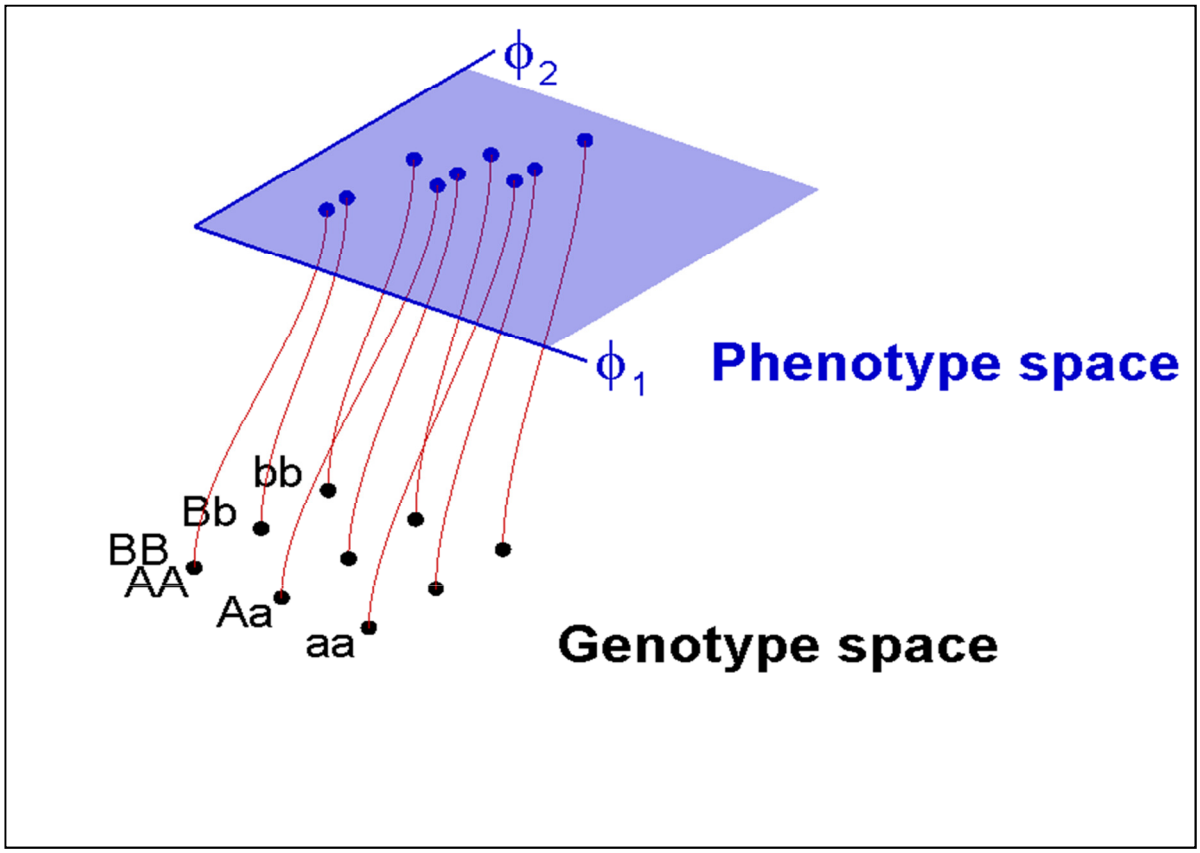


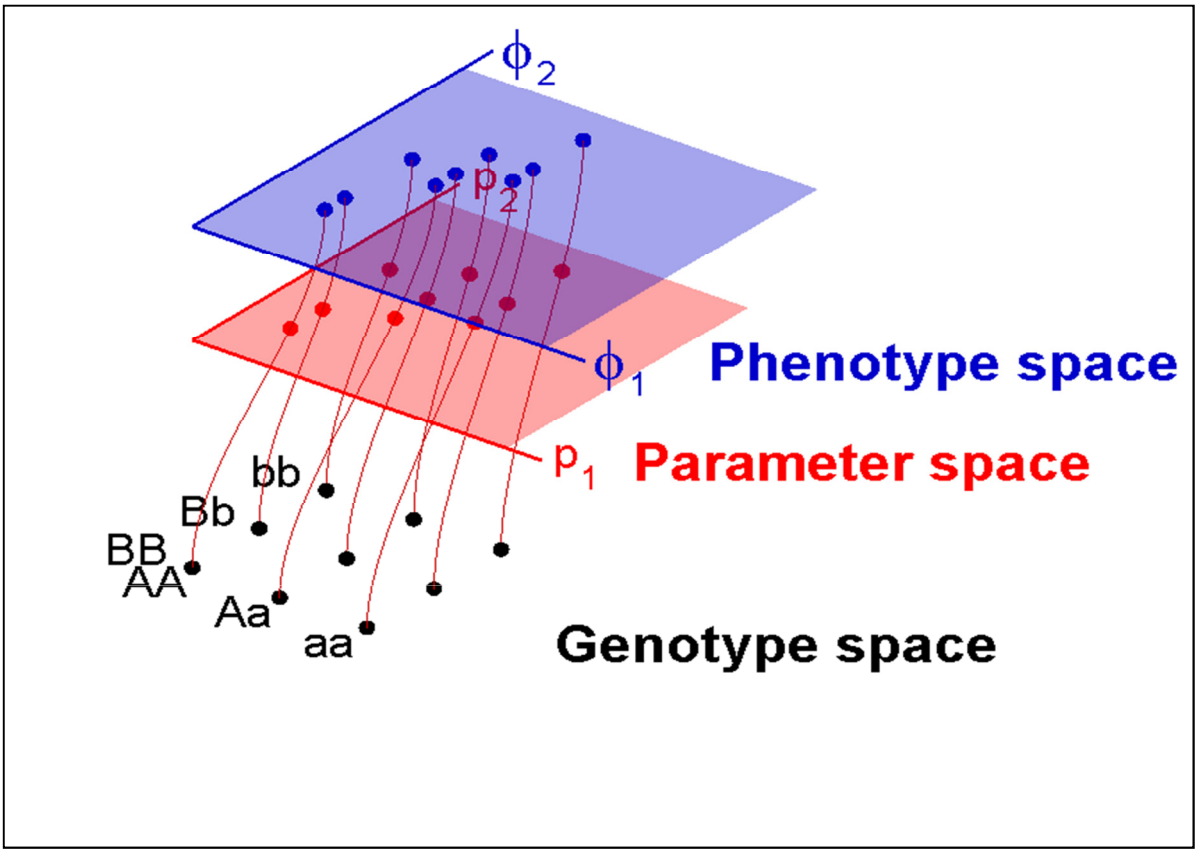
Here, the effect of a single copy of little-b is independent of the genotype at locus a, ...



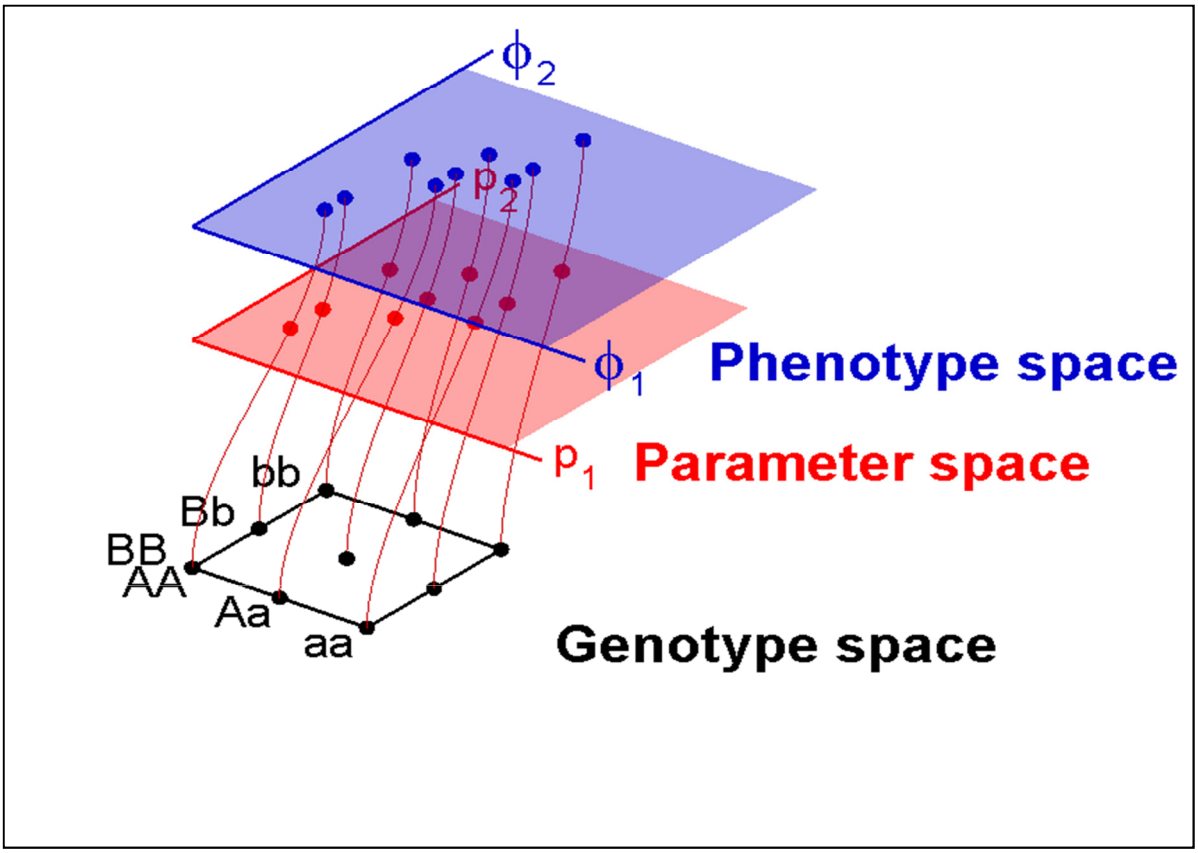
...but the effect of another "little b" depends on the genetic background, a phenomenon called *epistasis*.

Functional epistasis is here used as a common term for describing situations where the phenotypic effect of a genetic substitution (on one or multiple loci) depends on the genetic background, i.e. on the state of other loci in the genotype


ϕ_2
ϕ_1
Phenotype space
bb
Bb
BB
AA
Aa
aa
Genotype space



The mapping from genotypes to phenotypes can be decomposed into multiple transformations. The parameters of a physiological model can themselves be viewed as phenotypes, so that there is a mapping from genotype space to parameter space to phenotype space. Each of these can be further decomposed into layers of submodels.

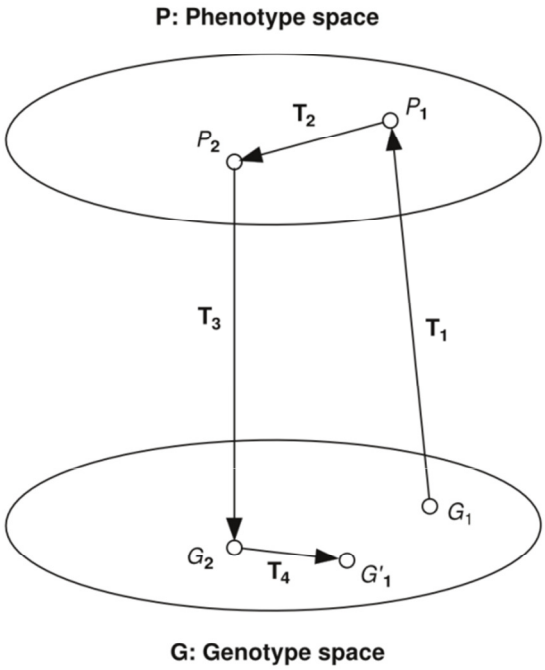


The mapping from genotype to parameters to phenotypes can be complex...

Introduce the causally cohesive genotype-phenotype model (cGP)

- cohesive: having the power of cohering, ie the power of sticking together

The GP map in evolution



R. C. Lewontin (1974): The genetic basis of evolutionary change

•The figure is a simplified version of the first figure in Lewontins excellent book from 1974

• Can be used to illustrate the current state of genetics theory as well as the genotype-phenotype map concept

• Describe the figure:

- Treat genotypes and phenotypes as state variables
- Study the transformations
  - T_1: the genotype to phenotype transition
  - T_2: natural selection, migration and mating
  - T_3: genotypes underlying selected phenotypes
  - T_4: the laws of recombination and segregation

• Population genetics theory is framed in genetic terms and make a caricature of the T1 mapping

• Quantitative genetics theory is framed in phenotypic terms and make a caricature of the T1 mapping

• Mark the T1 mapping and explain the GP map.

Bridging the gap: Modelling the causal chain from genes to phenotypes to population-level phenomena

- What Cigene is
- Bridging the gap: Overall goal
- Multilevel cGP models
  - The Physiome project
  - Multilevel heart model
- The single heart cell: A miniature multiscale system
  - Physiology and genetic basis
  - Mathematical model
  - Phenotypes
  - "Genotypic" parameter variation
- Preliminary results
  - Simulated phenotypic variation
  - Univariate and multivariate phenotypic distributions
  - Univariate and multivariate mathematical and statistical analyses

Context

- CIGENE, Centre for integrative genetics
  - Seek deep causal understanding of complex genetic traits
  - Systems-oriented computational biology
  - SNP genotyping platform
- Bridging the genotype-phenotype gap
  - eVITA project (NFR/RCN)
  - Small-scale detail: sequence data, molecular physiology
  - Large-scale patterns: statistical quantitative genetics and epidemiology
  - Little connection thus far
- Mammalian heart model:
  - causally cohesive genotype-phenotype (cGP) model
  - genetic → phenotypic variation at multiple levels (e.g. ion channel, cell, tissue, organ)
- Link to the IUPS Physiome Project

What we do

- Arne Gjuvsland, postdoc
  - Theoretical focus on genetics:
    G-P maps, cGP models vs. statistical quantitative genetics.
  - Study systems: Physiome heart model, yeast cell models
- Jon Olav Vik, postdoc
  - Analyzing multivariate data and behaviour of complex biological systems and cGP models. Main responsibility for the heart model work package.
- Øyvind Nordbø, PhD student
  - Conceptual focus on interface and transitions between multiple scales, and associated mathematical, numerical, and biological modelling challenges.
  - Continuum mechanics, biophysics. openCMISS software.
- Mary MacLachlan, postdoc
  - Modeling the ageing heart.
- Yunpeng Wang, PhD student
  - Using cGP models to resolve unaccounted-for heritability of complex disease.
- Stig W. Omholt, professor, group leader

Bridging the gap: Overall goal

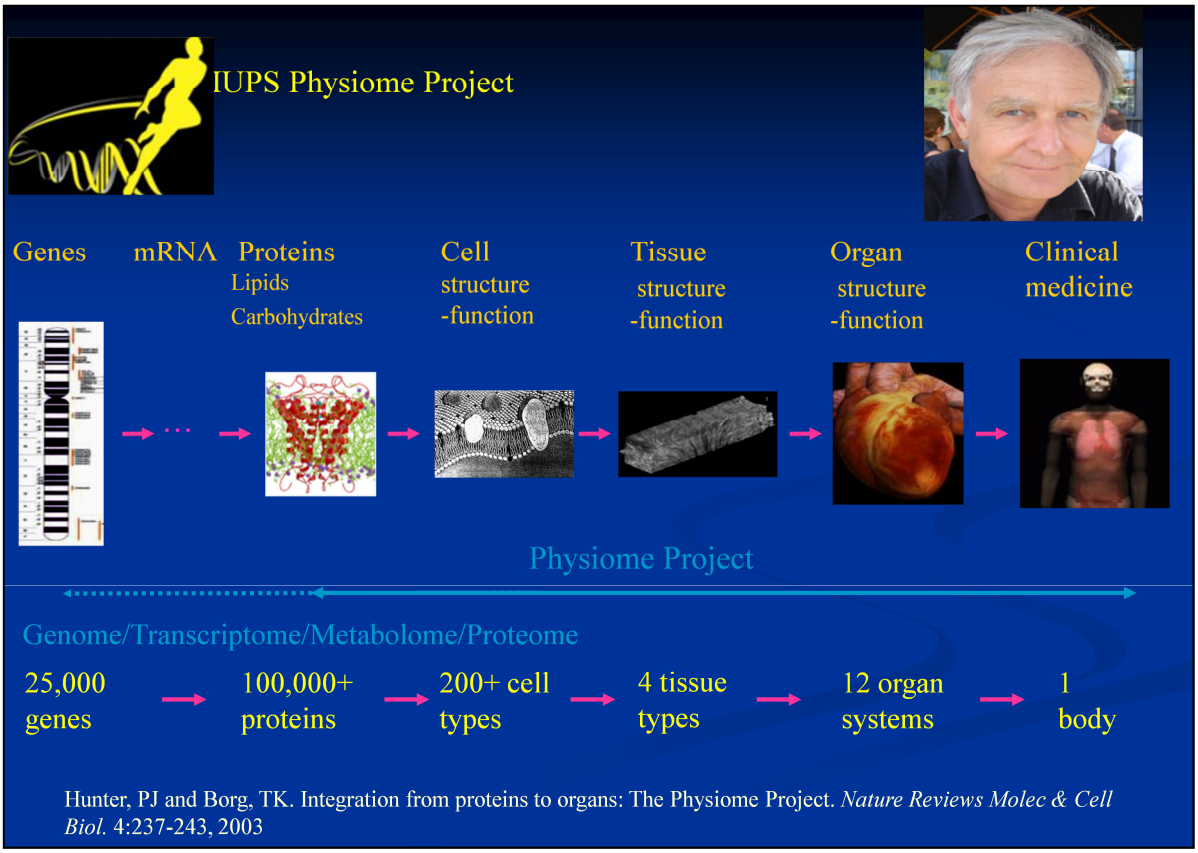
Big words

- Bridging the gap:
  disclosure, understanding and exploitation of the genotype-phenotype map
- *In a population setting establish a computational pipeline for describing, understanding and exploiting the link between genetic concepts and methodologies and the most advanced synthetic phenotype in the world, the mammalian heart model.*

- Analyze G-P map in existing single-cell models of mouse heart muscle cells
- Whole mouse heart modelling and analysis
- Analytical exercises related to model condensation and integration

Multilevel cGP models

- The Physiome project
- Multilevel heart model



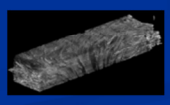
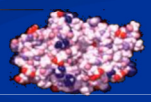
Forced to get the dimensions right.

The Challenge: spatial and temporal scales

Space 10^9
- 1 m person
- 1 mm electrical length scale of cardiac tissue
- 1 mm cardiac sarcomere spacing
- 1 nm pore diameter in a membrane protein

Time 10^{15}
- 10^9 s (70 yrs) human lifetime
- 10^6 s (10 days) protein turnover
- 10^3 s (1 hour) digest food
- 1 s heart beat
- 1 ms ion channel HH gating
- 1 ms Brownian motion



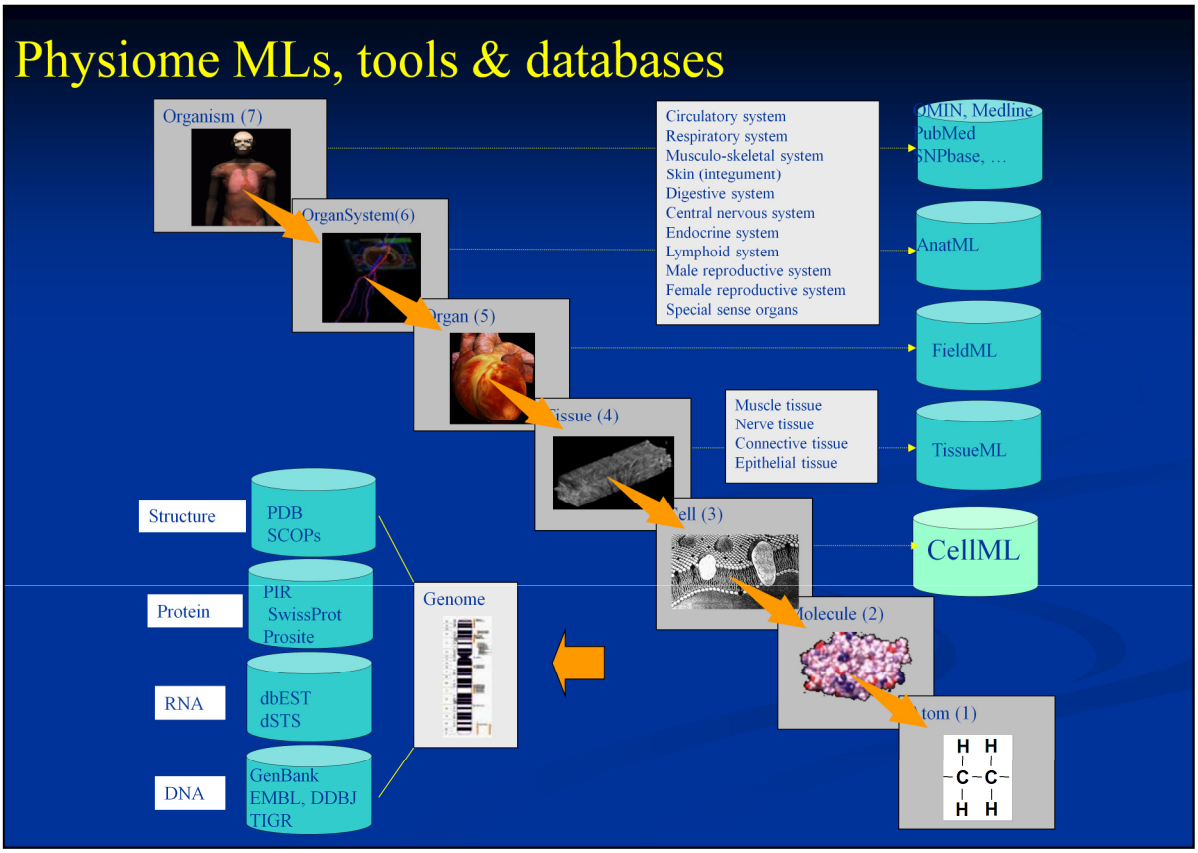




The diversity of experimental models

- bacterial models structural biology
- murine models functional genomics
- large animal models physiology
- human clinical MRI, CT, etc

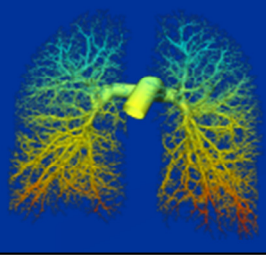
Requires a hierarchy of inter-related models

gene reg. networks → pathway models → MD/CG models → stochastic models → ODEs → PDEs (continuum models)


Physiome MLs, tools & databases
Organism (7)
OrganSystem(6)
Organ (5)
Tissue (4)
Cell (3)
Molecule (2)
Atom (1)
Circulatory system
Respiratory system
Musculo-skeletal system
Skin (integument)
Digestive system
Central nervous system
Endocrine system
Lymphoid system
Male reproductive system
Female reproductive system
Special sense organs
Muscle tissue
Nerve tissue
Connective tissue
Epithelial tissue
OMIN, Medline
PubMed
SNPbase, …
AnatML
FieldML
TissueML
CellML
Structure
PDB
SCOPs
Protein
PIR
SwissProt
Prosite
RNA
dbEST
dSTS
DNA
GenBank
EMBL, DDBJ
TIGR
Genome

Organ system Physiome Projects

Cardiovascular system
Respiratory system
Musculo-skeletal system
Digestive system
Skin (integument)
Urinary system
Lymphoid system
Female reproductive system
Special sense organs
Central nervous system
Endocrine system
Male reproductive system



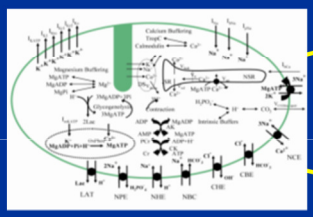
The Heart Physiome Project

torso

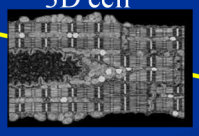
heart

tissue

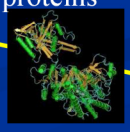
Cellular processes



cell-cell
connections



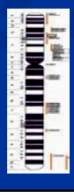
3D cell



proteins

transcription
& translation

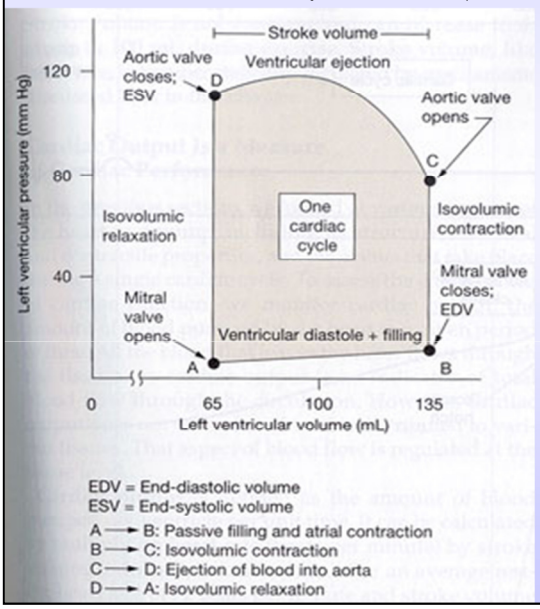
genes

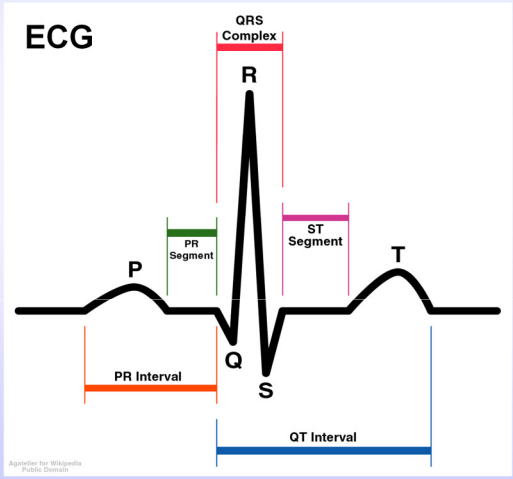


Hunter, PJ and Borg, TK. Integration from proteins to organs: The Physiome Project. *Nature Reviews Molec & Cell Biol.* 4:237-243, 2003

Phenotype measures

Cell phenotype	Tissue phenotype	Continuum models	Modes of deformation	Whole heart phenotype
Protein pathways	Spatial variation	Mechanics	Measures of arrhythmia	Heart rate
Electrophysiology	Fibre types	Electrical activation		Q-T interval
pH control	Fibre orientation	Ventricular blood flow		End diastolic volume
Signal transduction		Coronary flow		Left ventricular pressure
Calcium transport		Energetics		
Myofilament mechanics				





Phenotypes from cell to organ

Examples:

- Whole heart
  - Electrocardiogram
  - Arrhythmias
- Tissue
  - Activation propagation speed
  - Amplification or containment of beginning arrhythmia
- Cell
  - Action potential
  - Calcium dynamics
  - APD restitution
- Ion channel
  - Maximum conductance
  - Voltage dependence

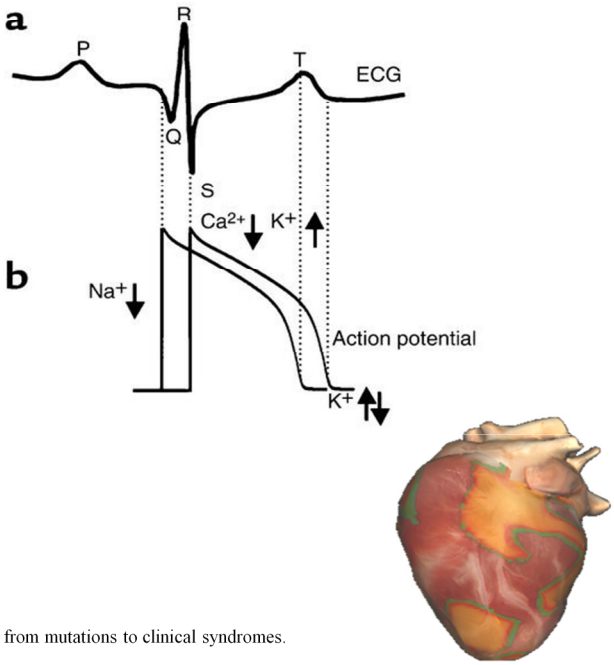


Figure: Clancy CE, Kass RS (2002) Defective cardiac ion channels: from mutations to clinical syndromes. J Clin Invest 110:1075–1077 http://dx.doi.org/10.1172/JCI0216945
Screenshot: http://thevirtualheart.org/vf.html

The cellular action potential underlies important organ-level phenotypes such as the electrocardiogram. The heartbeat starts in the atria (■first upstroke), causing a dip in the ECG. The so-called QRS complex follows as ventricular cells activate (■second upstroke). Repolarization (■ downstroke and T wave) marks the end of the action potential and the heartbeat.

A mutation which ■prolongs the action potential may cause ■"long QT syndrome", increasing the risk of lethal arrhythmias. This is not always the case, however. Whether cellular anomalies propagate to the whole-heart level depends on the spatial dynamics of electrical activation at the tissue level. Next year we will study the genotype-phenotype map in detailed models of cell tissue and the whole-heart. Finally, the penetrance of long QT seems to depend on both modifier genes and environmental effects.

Figure 1. Electrical gradients in the myocardium can be detected on the body surface ECG. (a) An illustrative example of a single cardiac cycle detected as spatial and temporal electrical gradients on the ECG. The P wave is generated by the spread of excitation through the atria. The QRS complex represents ventricular activation and is followed by the T wave reflecting ventricular repolarization gradients. (b) Schematic representation of cellular electrical activity underlying the ECG (see text for details). Where downward arrows represent inward current and upward arrows represent outward current.

The single heart cell: A miniature multiscale system

- Model: Li & Smith in prep (derived from Bondarenko et al. 2004)
  - 36 state variables (ordinary differential equations), 76 parameters
  - Fitted and validated against several experiments and cardiomyocyte types
  - Features in common with most single-cell models
- Similar challenges as whole-heart work
  - Summarize relevant features of complex phenotype
  - Statistical pattern vs actual mechanism
  - Sensitivity of phenotype to genotypic variation
  - Multivariate analyses: epistasis, pleiotropy, penetrance, expressivity
- Validate analysis methods
- Miniature multiscale
  - ion channels to ion concentrations to action potential

Realistic single-cell models can be quite complex genotype-phenotype maps in themselves.

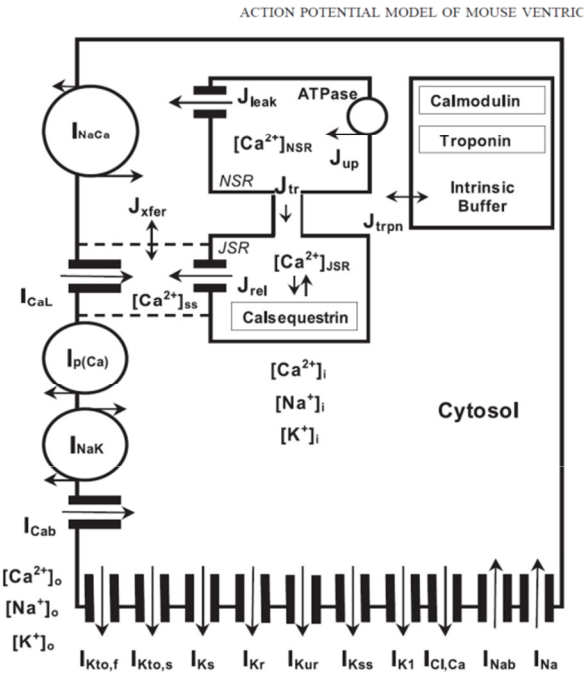
The parallels to our future work on the whole heart model in openCMISS is that we have a high dimensional phenotype that we need to summarize into something of clinical interest. There is also the duality of documenting the statistical pattern versus identifying actual mechanism. Two of the phenomena we have started looking at, are the effects of genotypic variation at one locus on the phenotype, and how such effects may modify the effect of variation at other loci. The latter phenomenon is called epistasis. Other phenomena of interest include pleiotropy, meaning that one gene affects several traits; penetrance, meaning that a genetic defect may not always show; and expressivity, meaning the degree of continuous variability in the effect of the genetic difference. This is a good opportunity to validate the analysis methods we are going to be using later, while at the same time producing real biological insights. It will also bolster our confidence when using these methods on whole heart simulations. Perhaps single cell models can even illustrate some features of multiscale systems. Their phenotypes certainly form the interface up to the tissue level.

The single heart cell: A miniature multiscale system

- Physiology and genetic basis
- Mathematical model
- Phenotypes
- "Genotypic" parameter variation

Model and parameters

- Heart muscle cell
  - Ion channels, exchangers and pumps, Ca^{2+} compartments
  - Reversal potentials, voltage sensitivity, …
  - Transitions between ion-channel configurations: Markov chains
  - Model of Bondarenko et al. (2004), refined by Li & Smith
- Low-level parameters
  - Maximum conductances for each current + Lots of others



(Figure 1 from Bondarenko et al. 2004, http://dx.doi.org/10.1152/ajpheart.00185.2003)

We have been studying a fairly detailed model of a mouse heart cell. It describes the flow of ions across the cell membrane and between different compartments of the cell. This flow of ions achieves the two main function of the cell: To contract the heart muscle, and to propagate an electrical signal.

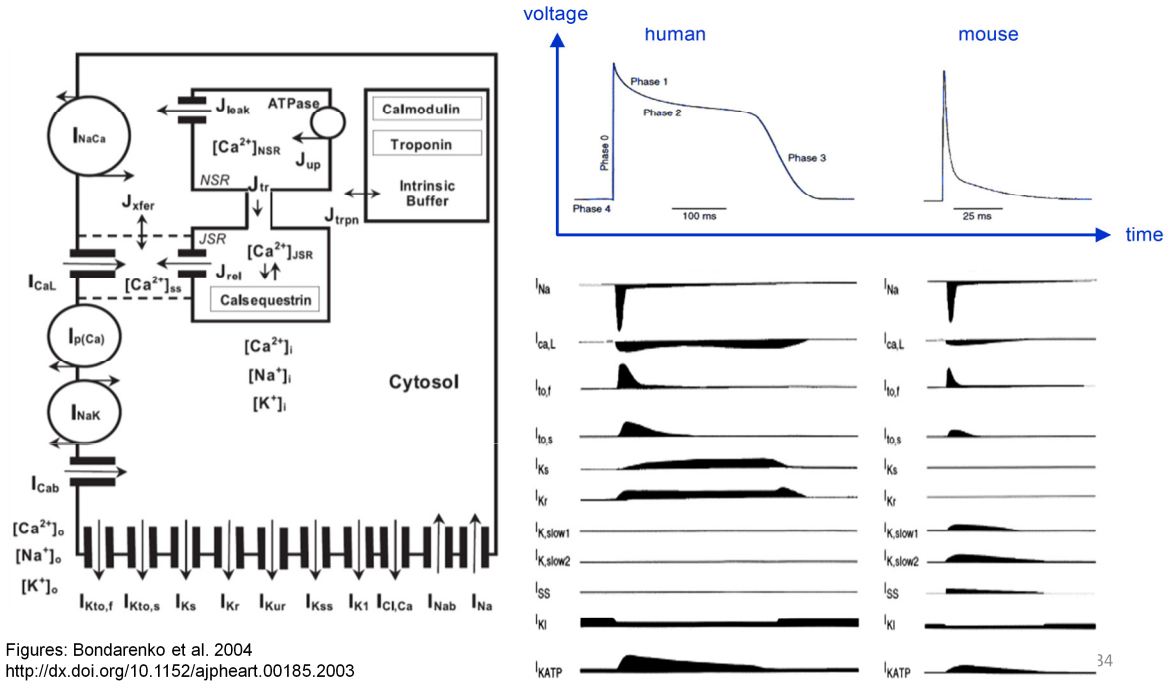
The cell "charges its battery" by moving many positive ions out of the cell fluid. Once it is charged, an electrical impulse will cause the cell to "fire", as ion channels open to allow rapid depolarization, producing a signal that propagates to neighbouring cells.

Muscle contraction, on the other hand, is triggered by the release of calcium into the cytosol. Initially, calcium is sequestered into special compartments, until depolarization triggers its release.

Now, I'm no expert on this, but I'll give it my best shot. The Bondarenko model describes the flow of ions in and out of the cell, and the resulting difference in electrical potential between the inside and the outside of the cell. This is called the transmembrane potential, and its time trajectory is what's called the action potential of the cell.

The major ions in the model are calcium, sodium, and potassium. Ion channels are specialized proteins that help or hinder the passage of ions across the cell membrane, opening or closing in response to conditions such as ion

Small-scale complexity: electrophysiology of a heart cell



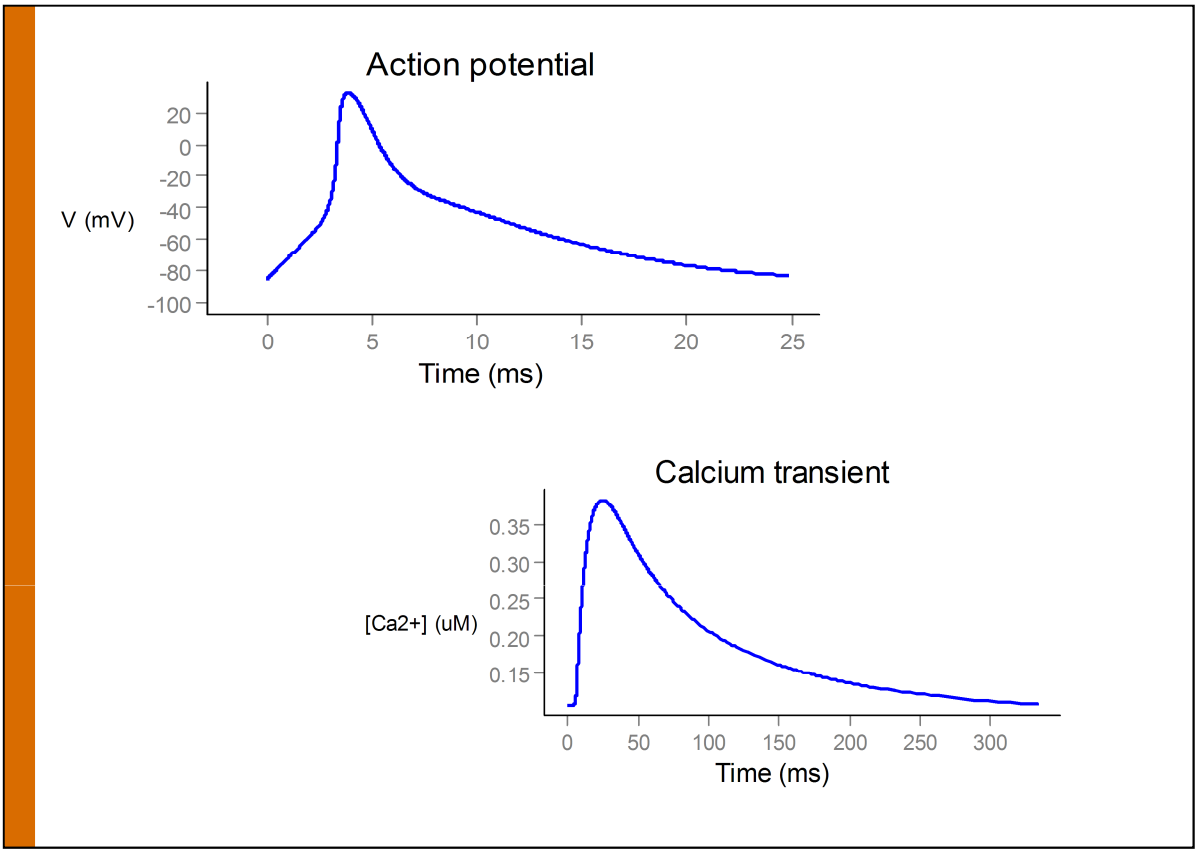
Figures: Bondarenko et al. 2004
http://dx.doi.org/10.1152/ajpheart.00185.2003
Nerbonne 2004 http://dx.doi.org/10.1016/j.tcm.2003.12.006

Complex gene interactions occur even within a single cell. Heart cells generate electrical impulses by varying the permeability of the ■cell membrane to ■sodium, potassium and calcium (Na+, K+, Ca2+) ions. Special proteins called ■ion channels open and close in response to changes in voltage or ion concentrations, determining whether ions are permitted to move along electrical or osmotic gradients. The complex interplay and timing of the various ion currents determine the time-course of the transmembrane voltage. This phenotype is so important that it has its own name: ■the action potential.

This is a compact example of multilevel phenotypes arising from complex gene interaction. ►■Genes code for different subunits which combine ■to form ion channels. Mutations may affect the conductance of ion channels, but also their opening and closing speed or voltage dependence. Multiple channels contribute to the ■concentration of each ion species, which in turn contribute to the ■transmembrane voltage. ►The resulting electrical impulse is the cell's primary interface to the surrounding tissue. This phenotype, the time-course of the transmembrane potential, is so important that it has its own name: the ■action potential.

Figure 1 from Bondarenko VE, Szigeti GP, Bett GCL, Kim S-J, Rasmusson RL (2004) Computer model of action potential of mouse ventricular myocytes. Am J Physiol Heart Circ Physiol 287(3):H1378–1403 http://dx.doi.org/10.1152/ajpheart.00185.2003

Figure 1 from Nerbonne JM (2004) Studying cardiac arrhythmias in the mouse-a reasonable model for probing mechanisms? Trends Cardiovasc Med 14(3):83–93 http://dx.doi.org/10.1016/j.tcm.2003.12.006



The time-course of the transmembrane voltage is called the *action potential*. The initial stimulus triggers rapid depolarization of the cell, and this signal is what propagates to neighbouring cells.

However, depolarization also triggers calcium release into the cytosol. This process is slower, as the muscle fibers need time to react. Meanwhile, the cell repolarizes, getting ready for the next action potential.

The action potential and calcium transient result from the combined action of many different ion channels, each coded for by specific genes.

Bondarenko model: ion channel Markov models

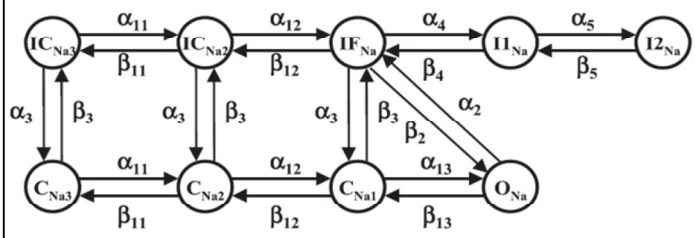


Fig. 2. State diagram of the Markov model for the Na^+ channel. C_{Na1}, C_{Na2}, and C_{Na3} are closed states; O_{Na} is the open state; IF_{Na} is the fast inactivated state; $I1_{Na}$ and $I2_{Na}$ are the intermediate inactivated states; and IC_{Na2} and IC_{Na3} are the closed-inactivation states (22). α and β are the transition rates between the states, as given in the APPENDIX.

AJP-Heart Circ Physiol • VOL 287 • SEPTEMBER 2004 • www.ajpheart.org

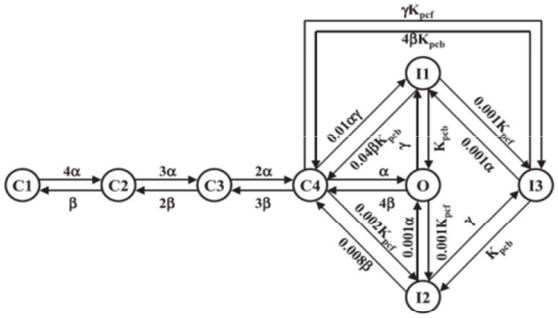


Fig. 4. State diagram of the model L-type Ca^{2+} channel. C_1, C_2, C_3, and C_4 are closed states; O is the open state; and I_1, I_2, and I_3 are inactivated states. The rate constants α and β are voltage dependent, and γ is calcium dependent. K_{pcf} and K_{pcb} are the forward and backward voltage-insensitive rate constants, respectively.

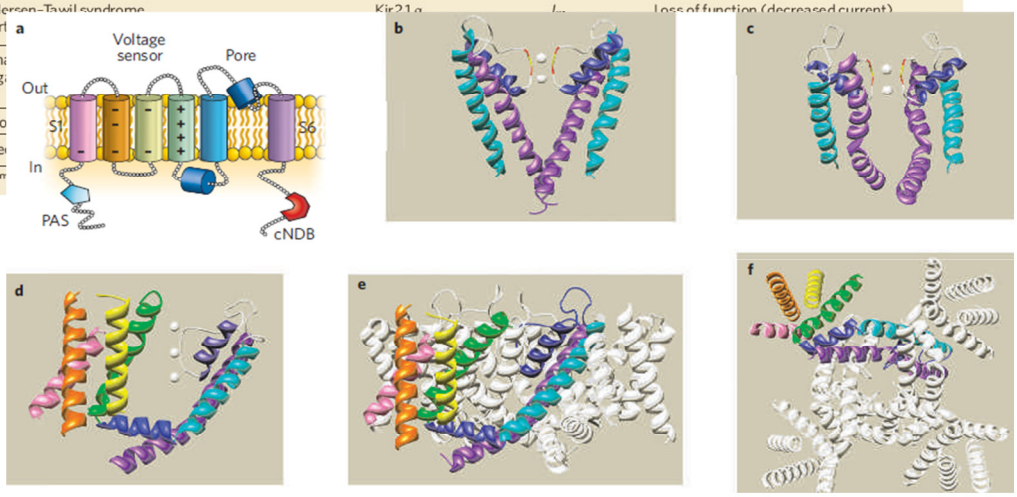
The major state variables in the model are the transmembrane potential, the ion concentrations, and the state of the ion channels. Ion channels are often modeled as being in one of several possible states, one of which is "open". A Markov chain model is a representation of the proportion of channels in each of these states and the transition rates between them. All of this may be an accurate representation of the number of subunits in the ion channel. But more often, it is just something to imitate the inertia and behaviour of the cannel proteins on average.

Ion channel genes

Table 1 | Ion-channel genes associated with cardiac arrhythmia

HERG[2,24] (*KCNH2*)	Romano–Ward (autosomal dominant) long QT syndrome Short QT syndrome	hERG (KV11.1)	I_{Kr}	Loss of function (decreased current) Gain of function (impaired inactivation)
KCNQ1[24,81]	Romano–Ward and Jervell and Lange-Nielsen (autosomal recessive) long QT syndromes Short QT syndrome	KvLQT1 (KV7.1) α	I_{Ks}	Loss of function (decreased current) Gain of function (channels remain open)
KCNE1[82]	Romano–Ward & Jervell and Lange-Nielsen long QT syndromes	minK β	I_{Ks}	Loss of function (decreased current)
KCNE2[38]	Romano-Ward long QT syndromes	MiRP1 β	I_{Kr}	Loss of function (decreased current)
KCNJ2[16,87]	Short			
SCN5A[83,84,88]	Roma Brug			
CACNA1C[17]	Timo			
RYR2[85]	Cate			

Alternate gene or subunit nam



Several hundred known ion-channel mutations are associated with cardiac arrhythmia, and their function is beginning to be well understood. This (lower right) is looking out through an ion channel consisting of four identical subunits. Each subunit has a voltage-sensing part with positively charged amino acid residues. The sensor slides across the cell membrane in response to voltage changes, pushing on a part of the pore to close the channel.

Bondarenko model: parameters and equations

Table 3. *Extracellular ion concentrations*

Parameter	Definition	Value, μM
$[K^+]_o$	Extracellular K^+ concentration	5,400
$[Na^+]_o$	Extracellular Na^+ concentration	140,000
$[Ca^{2+}]_o$	Extracellular Ca^{2+} concentration	1,800

$$\frac{d[Ca^{2+}]_{ss}}{dt} = B_{ss}\left\{J_{rel}\frac{V_{JSR}}{V_{ss}} - J_{xfer}\frac{V_{myo}}{V_{ss}} - I_{CaL}\frac{A_{cap}C_m}{2V_{ss}F}\right\} \quad (A3)$$

$$\frac{d[Ca^{2+}]_{JSR}}{dt} = B_{JSR}\{J_{tr} - J_{rel}\} \quad (A4)$$

$$\frac{d[Ca^{2+}]_{NSR}}{dt} = \{J_{up} - J_{leak}\}\frac{V_{myo}}{V_{NSR}} - J_{tr}\frac{V_{JSR}}{V_{NSR}} \quad (A5)$$

$$B_i = \left\{1 + \frac{[CMDN]_{tot}K_m^{CMDN}}{(K_m^{CMDN} + [Ca^{2+}]_i)^2}\right\}^{-1} \quad (A6)$$

$$B_{ss} = \left\{1 + \frac{[CMDN]_{tot}K_m^{CMDN}}{(K_m^{CMDN} + [Ca^{2+}]_{ss})^2}\right\}^{-1} \quad (A7)$$

$$B_{JSR} = \left\{1 + \frac{[CSQN]_{tot}K_m^{CSQN}}{(K_m^{CSQN} + [Ca^{2+}]_{JSR})^2}\right\}^{-1} \quad (A8)$$

Calcium fluxes.

$$J_{rel} = v_1(P_{O1} + P_{O2})([Ca^{2+}]_{JSR} - [Ca^{2+}]_{ss})P_{RyR} \quad (A9)$$

$$J_{tr} = \frac{[Ca^{2+}]_{NSR} - [Ca^{2+}]_{JSR}}{\tau_{tr}} \quad (A10)$$

$$J_{xfer} = \frac{[Ca^{2+}]_{ss} - [Ca^{2+}]_i}{\tau_{xfer}} \quad (A11)$$

$$J_{leak} = v_2([Ca^{2+}]_{NSR} - [Ca^{2+}]_i) \quad (A12)$$

$$J_{up} = v_3\frac{[Ca^{2+}]_i^2}{K_{m,up}^2 + [Ca^{2+}]_i^2} \quad (A13)$$

Table 5. *L-type Ca^{2+} channel parameters*

Parameter	Definition	Value
G_{CaL}	Specific maximum conductivity for L-type Ca^{2+} channel	0.1729 mS/μF
$E_{Ca,L}$	Reversal potential for L-type Ca^{2+} channel	63.0 mV
$K_{pc,max}$	Maximum time constant for Ca^{2+}-induced inactivation	0.23324 ms^{-1}
$K_{pc,half}$	Half-saturation constant for Ca^{2+}-induced inactivation	20.0 μM
K_{pcb}	Voltage-insensitive rate constant for inactivation	0.0005 ms^{-1}
$I_{CaL,max}$	Normalization constant for L-type Ca^{2+} current	7.0 pA/pF

$$J_{trpn} = k_{htrpn}^+[Ca^{2+}]_i([HTRPN]_{tot} - [HTRPNCa]) - k_{htrpn}^-[HTRPNCa] + k_{ltrpn}^+[Ca^{2+}]_i([LTRPN]_{tot} - [LTRPNCa]) - k_{ltrpn}^-[LTRPNCa] \quad (A14)$$

$$\frac{dP_{RyR}}{dt} = -0.04P_{RyR} - 0.1\frac{I_{CaL}}{I_{CaL,max}}e^{\frac{-(V-5.0)^2}{648.0}} \quad (A15)$$

Calcium buffering.

$$\frac{d[LTRPNCa]}{dt} = k_{ltrpn}^+[Ca^{2+}]_i([LTRPN]_{tot} - [LTRPNCa]) - k_{ltrpn}^-[LTRPNCa] \quad (A16)$$

$$\frac{d[HTRPNCa]}{dt} = k_{htrpn}^+[Ca^{2+}]_i([HTRPN]_{tot} - [HTRPNCa]) - k_{htrpn}^-[HTRPNCa] \quad (A17)$$

Ryanodine receptors.

$$\frac{dP_{O1}}{dt} = k_a^+[Ca^{2+}]_{ss}^2P_{C1} - k_a^-P_{O1} - k_b^+[Ca^{2+}]_{ss}^3P_{O1} + k_b^-P_{O2} - k_c^+P_{O1} + k_c^-P_{C2} \quad (A18)$$

$$P_{C1} = 1 - (P_{C2} + P_{O1} + P_{O2}) \quad (A19)$$

$$\frac{dP_{O2}}{dt} = k_b^+[Ca^{2+}]_{ss}^4P_{O1} - k_b^-P_{O2} \quad (A20)$$

$$\frac{dP_{C2}}{dt} = k_c^+P_{O1} - k_c^-P_{C2} \quad (A21)$$

This is just part of the listing of the model parameters and equations of the Bondarenko model.

Bondarenko model: parameters and initial conditions

Table 7. *Membrane current parameters*

Parameter	Definition	Value
C_m	Specific membrane capacitance	1.0 μF/cm²
F	Faraday constant	96.5 C/mmol
T	Absolute temperature	298 K
R	Ideal gas constant	8.314 J·mol⁻¹·K⁻¹
k_{NaCa}	Scaling factor of Na^+/Ca^{2+} exchange	292.8 pA/pF
$K_{m,Na}$	Na^+ half-saturation constant for Na^+/Ca^{2+} exchange	87,500 μM
$K_{m,Ca}$	Ca^{2+} half-saturation constant for Na^+/Ca^{2+} exchange	1,380 μM
k_{sat}	Na^+/Ca^{2+} exchange saturation factor at very negative potentials	0.1
η	Controls voltage dependence of Na^+/Ca^{2+} exchange	0.35
I_{NaK}^{max}	Maximum Na^+/K^+ exchange current	0.88 pA/pF
$K_{m,Nai}$	Na^+ half-saturation constant for Na^+/K^+ exchange current	21,000 μM
$K_{m,Ko}$	K^+ half-saturation constant for Na^+/K^+ exchange current	1,500 μM
$I_{p(Ca)}^{max}$	Maximum Ca^{2+} pump current	1.0 pA/pF
$K_{m,p(Ca)}$	Ca^{2+} half-saturation constant for Ca^{2+} pump current	0.5 μM
G_{Cab}	Maximum background Ca^{2+} current conductance	0.000367 mS/μF
G_{Na}	Maximum fast Na^+ current conductance	13.0 mS/μF
G_{Nab}	Maximum background Na^{2+} current conductance	0.0026 mS/μF
$G_{Kto,f}$	Maximum transient outward K^+ current conductance (apex)	0.4067 mS/μF
$G_{Kto,\ f}$	Maximum transient outward K^+ current conductance (septum)	0.0798 mS/μF
G_{Ks}	Maximum slow delayed-rectifier K^+ current conductance	0.00575 mS/μF
$G_{Kto,s}$	Maximum transient outward K^+ current conductance (apex)	0.0 mS/μF

$$\frac{dC_3}{dt} = 3\alpha C_2 - 2\beta C_3 + 3\beta C_4 - 2\alpha C_3 \qquad (A26)$$

$$\frac{dC_4}{dt} = 2\alpha C_3 - 3\beta C_4 + 4\beta O - \alpha C_4 + 0.01(4K_{pcb}\,\beta I_1 - \alpha\gamma C_4) + 0.002(4\beta I_2 - K_{pcf}\,C_4) + 4\beta K_{pcb}\,I_3 - \gamma K_{pcf}\,C_4 \qquad (A27)$$

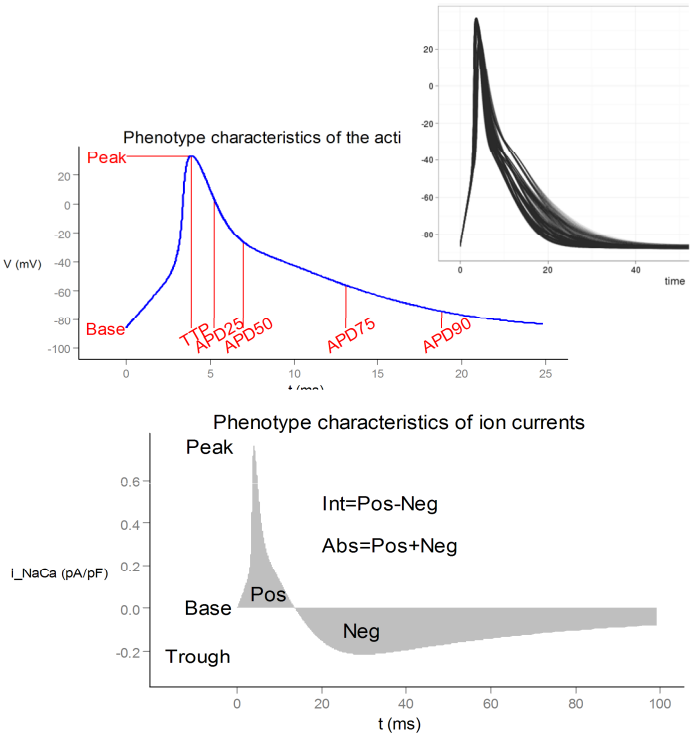
Table 8. *Initial conditions*

Parameter	Definition	Value
t	Time	0.0 ms
V	Membrane potential	−82.4202 mV
$[Ca^{2+}]_i$	Myoplasmic Ca^{2+} concentration	0.115001 μM
$[Ca^{2+}]_{ss}$	Subspace SR Ca^{2+} concentration	0.115001 μM
$[Ca^{2+}]_{JSR}$	JSR Ca^{2+} concentration	1299.50 μM
$[Ca^{2+}]_{NSR}$	NSR Ca^{2+} concentration	1299.50 μM
[LTRPNCa]	Concentration Ca^{2+} bound low-affinity troponin-binding sites	11.2684 μM
[HTRPNCa]	Concentration Ca^{2+} bound high-affinity troponin-binding sites	125.290 μM
O	L-type Ca^{2+} channel conducting state	0.930308×10^{-18}
C_1	L-type Ca^{2+} channel closed state	0.999876
C_2	L-type Ca^{2+} channel closed state	0.124216×10^{-3}
C_3	L-type Ca^{2+} channel closed state	0.578679×10^{-8}
C_4	L-type Ca^{2+} channel closed state	0.119816×10^{-12}
I_1	L-type Ca^{2+} channel inactivated state	0.497923×10^{-18}
I_2	L-type Ca^{2+} channel inactivated state	0.345847×10^{-13}
I_3	L-type Ca^{2+} channel inactivated state	0.185106×10^{-13}
P_{C1}	Fraction of RyR channels in state P_{C1}	0.999817
P_{C2}	Fraction of RyR channels in state P_{C2}	0.167740×10^{-3}
P_{O1}	Fraction of RyR channels in state P_{O1}	0.149102×10^{-4}
P_{O2}	Fraction of RyR channels in state P_{O2}	0.951726×10^{-10}
P_{RyR}	RyR modulation factor	0.0
C_{Na3}	Closed state of fast Na^+ channel	0.624646
C_{Na2}	Closed state of fast Na^+ channel	0.020752

This graph illustrates how we can summarize the action potential (the time trajectory) in a few key numbers: the resting and peak potentials, and the time required to get half way back down. The phenotype depends on parameter values, as illustrated here for cells from two different parts of the heart. One is the apex (the tip of the heart); the other is the septum (the dividing wall). Assuming that specific loci code for specific parameters, we may simulate phenotypic variation reflecting genetic variation, and the statistics on it. On the other hand, we can detail the mechanism by which the statistical pattern arises.

"Genotypes" and Phenotypes

- "Genotypic" variation
  - Maximum conductances of eight ion currents
  - "aa", "Aa", "AA" = 50%, 100%, 150%
  - 3^n parameter scenarios
- Phenotypes
  - Action potential (time-course of transmembrane voltage)
  - Ca^{2+} transient (contraction coupling)
  - Damping ratio of ion channels
  - Combined current of each ion species

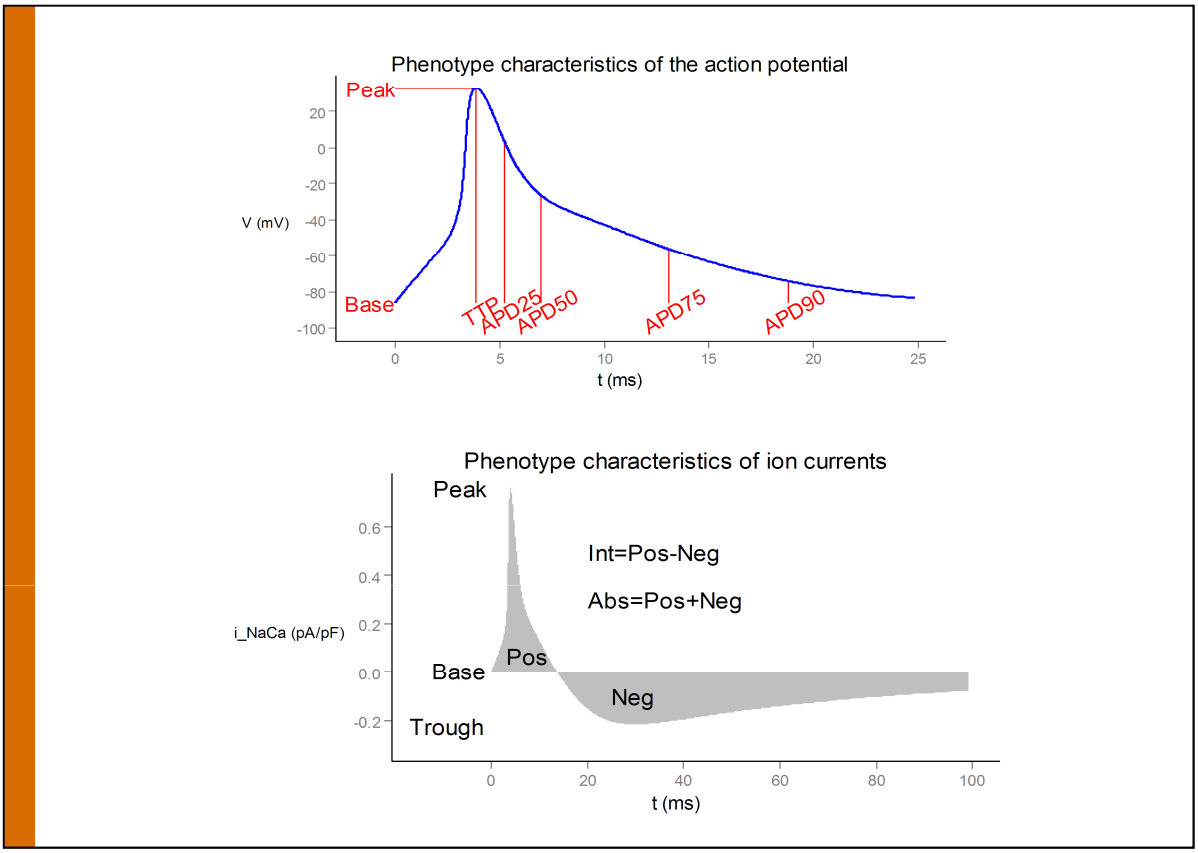


Many genetic analyses require that the high-dimensional phenotype of a cGP model be summarized in a few key statistics, such as the action potential duration to 90% repolarization. Similar measures are APD75, APD50, and APD25, as well as the time to peak, and peak voltage.

(Maybe skip this: At the subcellular level, ion-channel phenotypes can be described by total current during the action potential, or the timing and amplitude of the peak and trough.)

The next slide shows the distribution of each of these phenotypes, for six and a half thousand parameter sets resulting from "genetic" variation in eight parameters.

(This model is based on the Bondarenko 2004 model with modifications to the Cai dynamics. Specifically, the L-type Ca current, SERCA, NCX and Ryanodine receptor have been fitted to experimental data from our collaborators, and parameters for the background Ca current and PMCA have been adjusted accordingly.)



Many genetic analyses require that the high-dimensional phenotype of a cGP model be summarized in a few key statistics, such as the action potential duration to 90% repolarization. Similar measures are APD75, APD50, and APD25, as well as the time to peak, and peak voltage.

(Maybe skip this: At the subcellular level, ion-channel phenotypes can be described by total current during the action potential, or the timing and amplitude of the peak and trough.)

The next slide shows the distribution of each of these phenotypes, for six and a half thousand parameter sets resulting from "genetic" variation in eight parameters.

Expectations/Hypotheses

- Additive action for phenotypes relating mainly to a single current, e.g. upstroke (peak amplitude, time to peak) is mainly due to INa
- Effects of one gene may change the playing field for another:
  - e.g. if gene A affects voltage or Ca^{2+}, and gene B codes for a voltage- or Ca^{2+} -sensitive mechanism
- Some very obvious ones
  - SERCA is the Bondarenko model's only inflow of Ca^{2+} to the sarcoplasmic reticulum

42

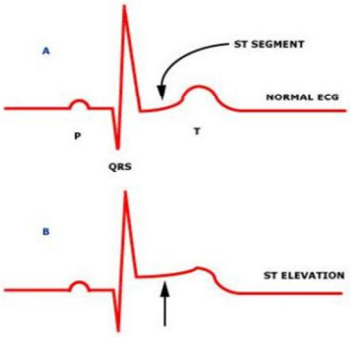
Epistasis, and measures thereof

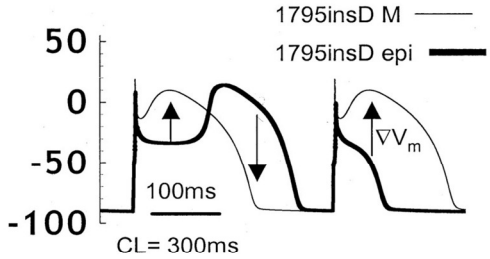
- Basic idea: The genotype at one locus modifies the effect of a gene substitution at another locus (for a given phenotype).
- Depends on the choice of phenotype and how much alleles vary.
- A kind of parameter interaction. Possible measures:
  - Second derivatives: $\partial^2(\text{pheno})/[\partial(\text{param } i)\ \partial(\text{param } j)]$
  - Regression coefficients for interaction terms in statistical models, for a given coding of genotypes. Appearance of epistasis depends on coding and genotype frequencies.
  - NOIA: "natural and orthogonal interactions". Can translate between different genotype frequencies.

43

Gene × Cell-type interaction

- Mutation *SCN5a*^{1795insD} causes ST elevation, a symptom of Brugada syndrome.
- Alternating "coved dome" and loss of plateau in epicardial cells, but not in midmyocardium (M cells)
- The affected ion channel is less important in M cells





Clancy CE, Rudy Y (2002) Na+ channel mutation that causes both Brugada and long-QT syndrome phenotypes. Circulation 105:1208–1213 http://dx.doi.org/10.1161/hc1002.105183
http://www.uptodate.com/online/content/image.do?imageKey=card_pix/st_patte.htm

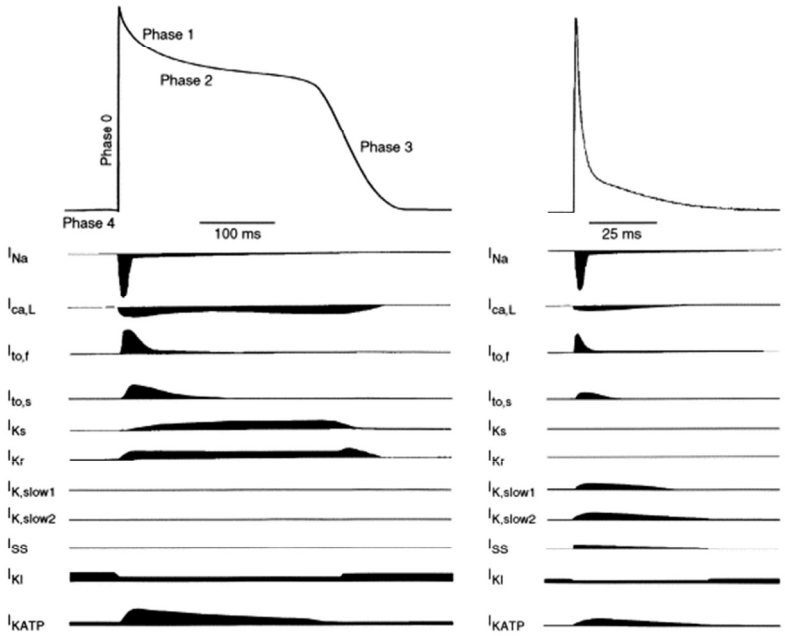
The effect of a mutation may differ between cell types. This mutation is linked to Brugada syndrome, which manifests as elevation of the ST segment of the ECG. However, the action potential is roughly normal in mutant midmyocardial cells (thin line), whereas epicardial cells show dramatic alternation between delayed and near-absent plateau phases of the action potential. The reason is that the affected ion channel is less expressed in M cells.

Figure 8. Effects of1795insD on transmural voltage gradients and APD. A, At fast rates, mutation-induced changes in epicardial AP morphologies (thick line) cause dispersion of plateau potentials and a voltage gradient (Vm, arrows) from epicardial to M cell (thin line). This gradient will manifest on the ECG as ST-segment elevation, indicative of Brugada syndrome. For coved-dome morphology of the epicardial AP, Vm is reversed during phase 3 repolarization, which can cause T-wave inversion on the ECG. B, At a slow rate (CL=850 ms), mutation prolongs APD in M cells (thick line) compared with WT (thin line). Delay in repolarization (APD60 ms) is reflected as QT prolongation on the ECG, a hallmark of LQT syndrome.

Preliminary results

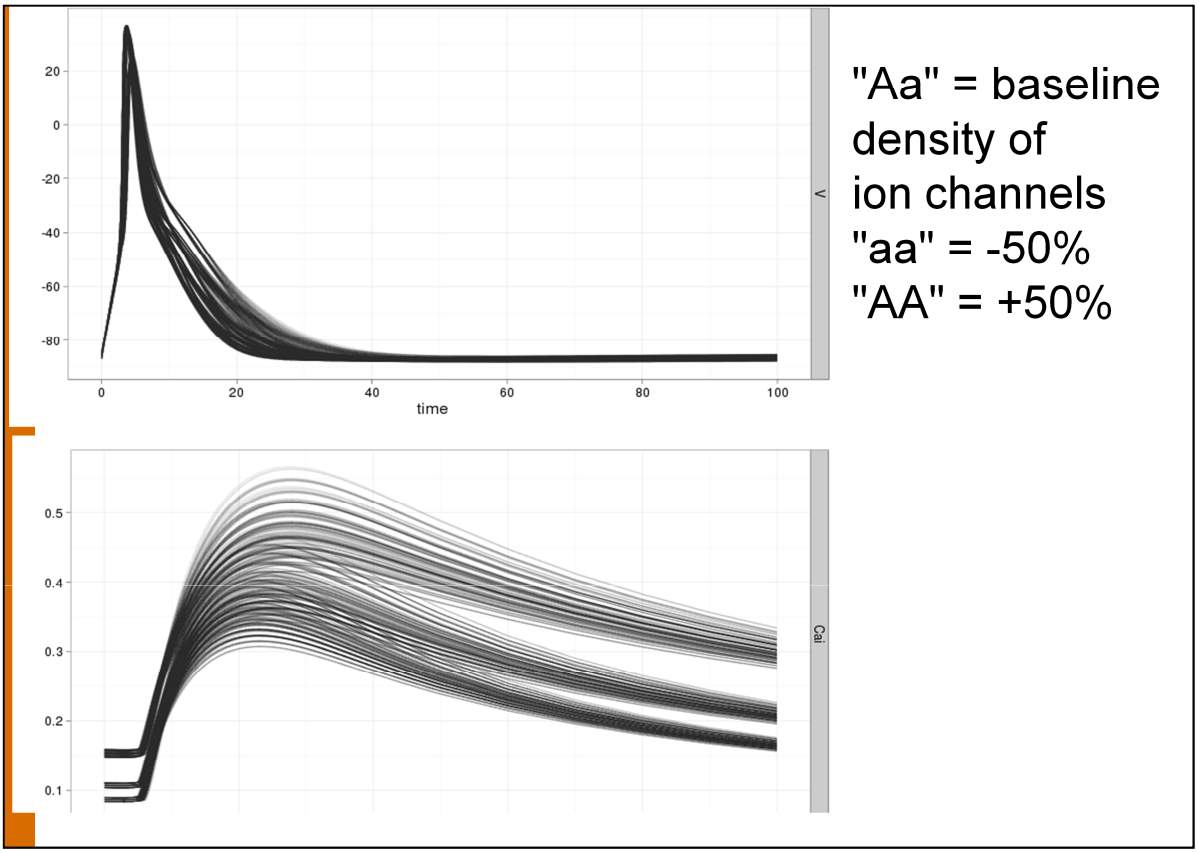
- Simulated phenotypic variation
- Univariate and multivariate phenotypic distributions
- Univariate and multivariate mathematical and statistical analyses

Ion currents in humans and mice



Different ion channels are responsible for different parts of the action potential. The upstroke is mainly due to the sodium current, whereas various potassium currents are important in repolarization.

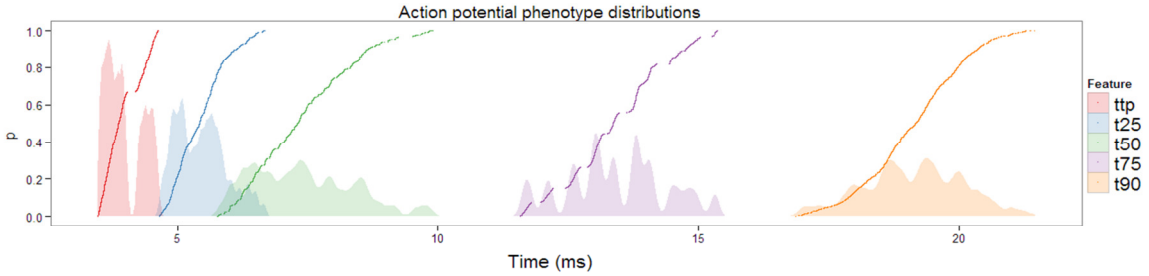
Now, a mutation may change the biochemical properties of an ion channel. When model parameters change, the phenotypes change too...



...Our aim to to understand how and why phenotypes vary in response to genetic variation in low-level parameters.

For simplicity, we assume that eight "genes" code for one ion channel each. For each gene, genotype "big A, little a" codes for the normal amount of ion channels, whereas one homozygote has 50 percent less and the other has 50 percent more. Variation in one gene then gives three different parameter scenarios; variation in two genes gives nine scenarios; and so on with successive splitting in threes. For example, the calcium transient shows a strong effect of a single gene (non-additive, it seems), and minor effects of other genes.

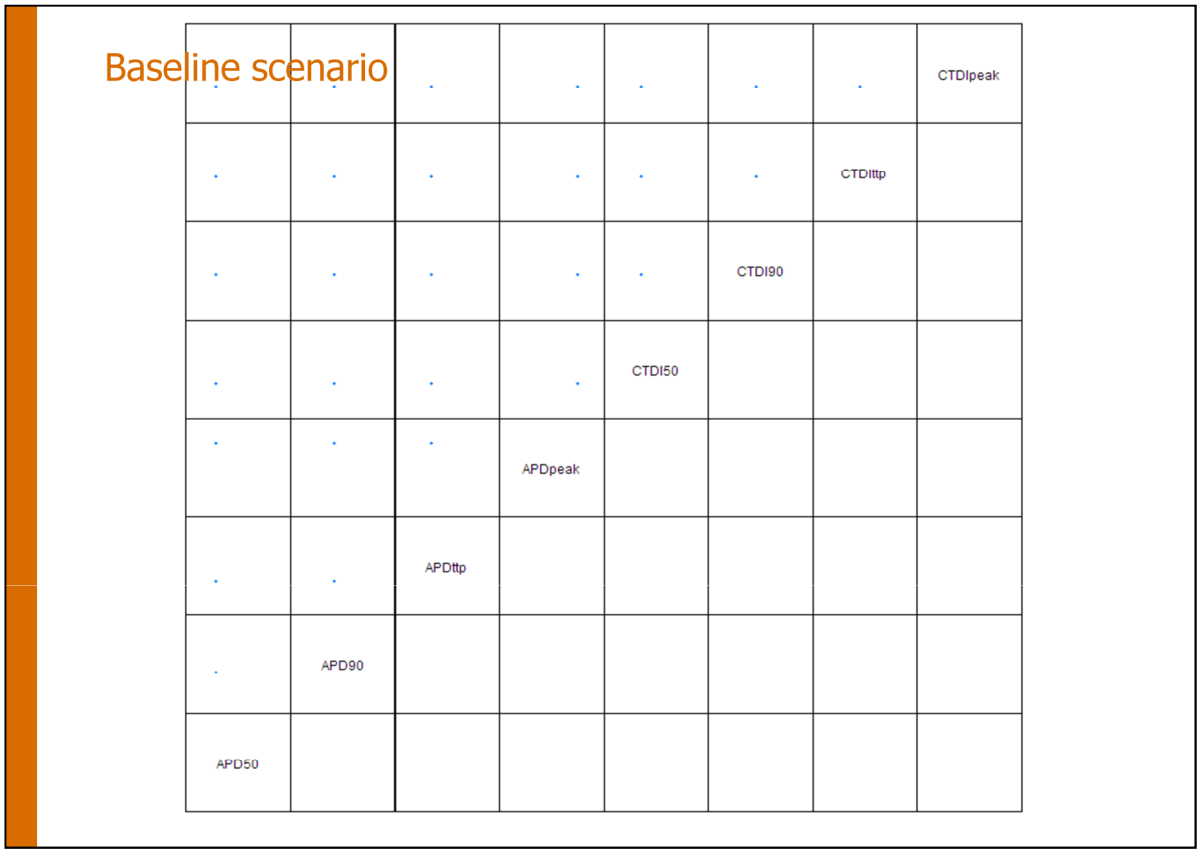
Univariate phenotype distributions



It is interesting to see how phenotypes differ in the number of underlying genes. The time to 90% repolarization, called APD90, is a fairly polygenic trait, showing quite continuous variation. In contrast, APD75 has a nine-modal distribution suggesting two underlying genes. And the time-to-peak phenotype appears to be governed largely by a single gene.

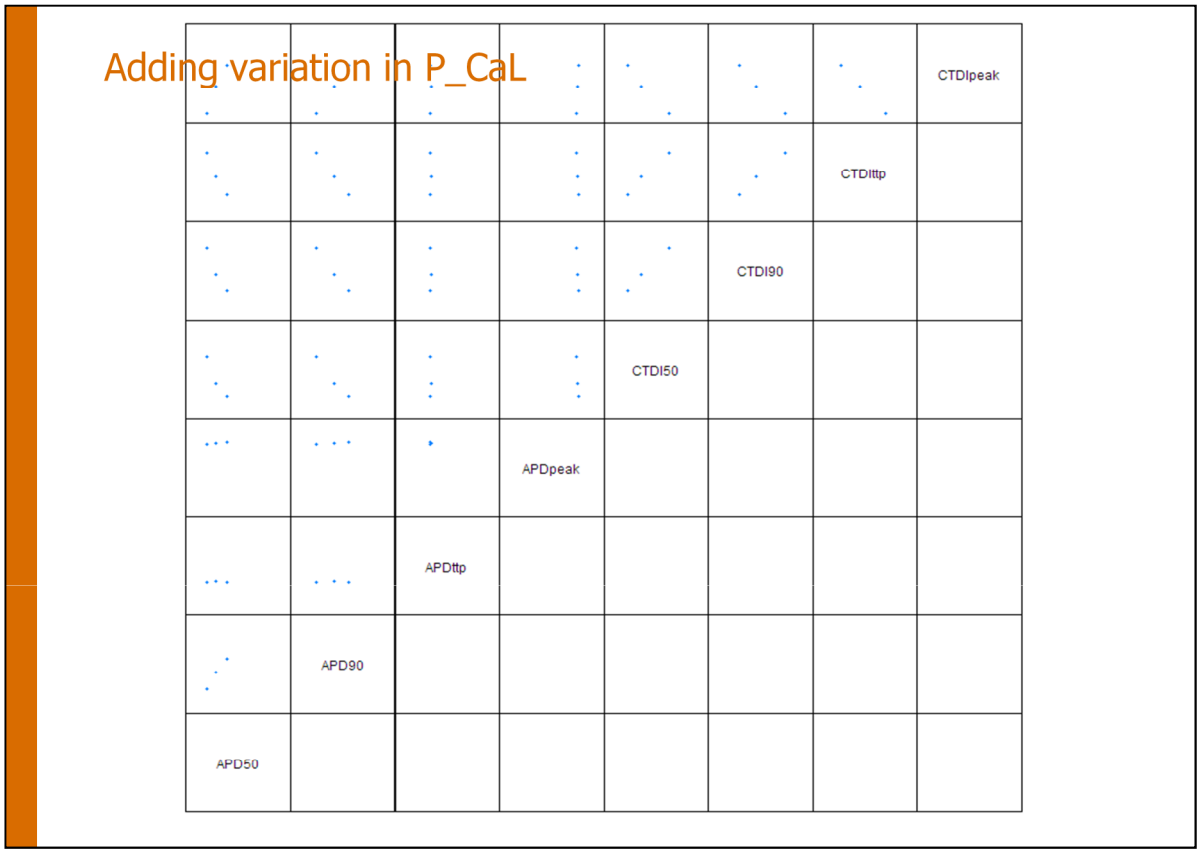
[The partly broken lines are empirical cumulative distribution functions, ECDF. Areas show kernel density estimates.]

Splitting 3x3x3...

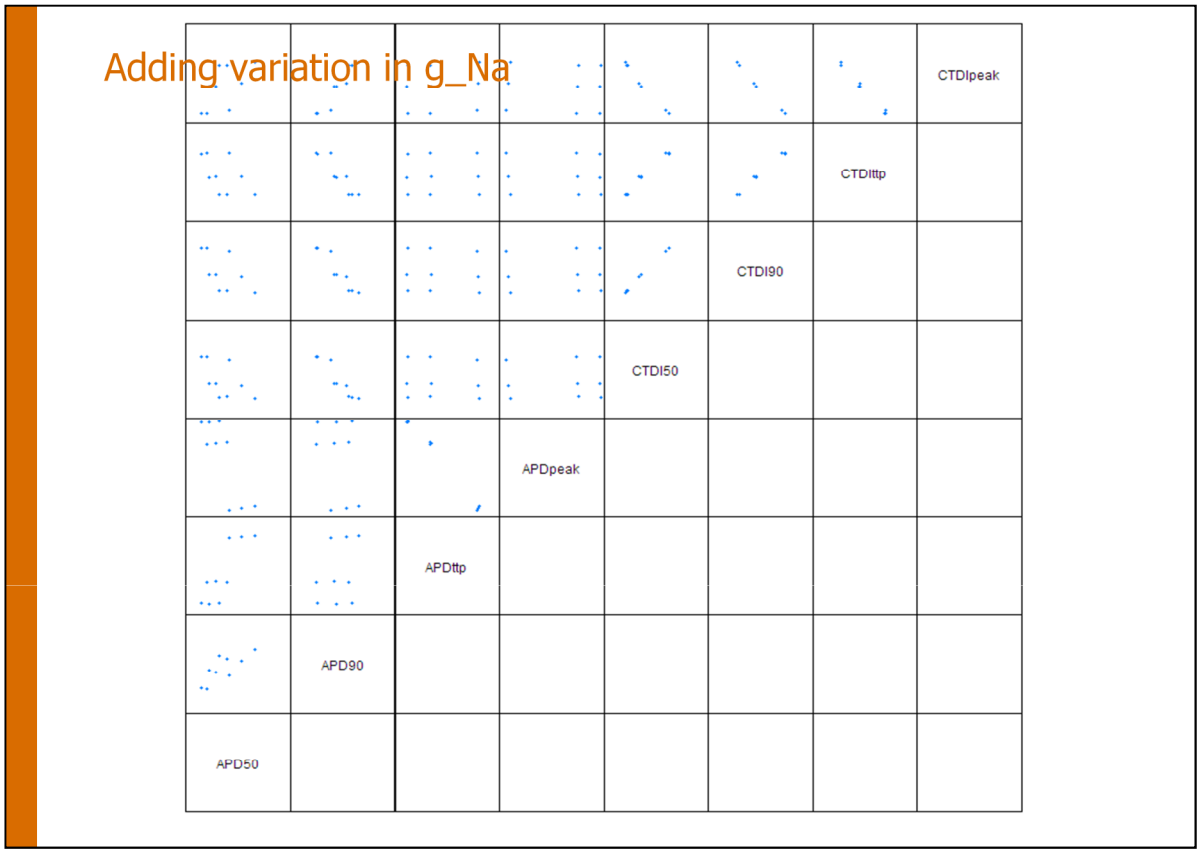


This is a scatterplot matrix with one panel for each pair of phenotypes. Currently, we see the phenotypes for the baseline parameter scenario.

(APD90 = action potential duration to 90% repolarization. "ttp" = time to peak. "peak" = peak value. CTDi = calcium transient duration phenotypes (the "i"means in cytosol) analogous to APD.)

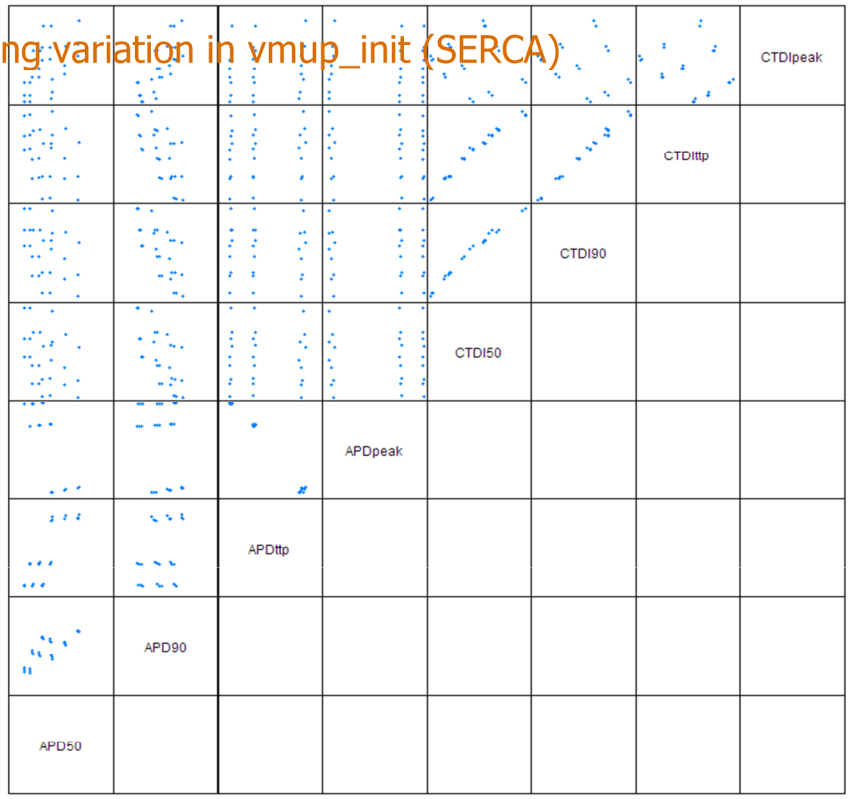


Varying parameter P_CaL (which sets the maximum of the L-type calcium current) gives us three parameter scenarios. APDttp and APDpeak are not much affected by this variation (since they occur earlier than the calcium transient), whereas the other phenotypes do respond to this parameter variation.



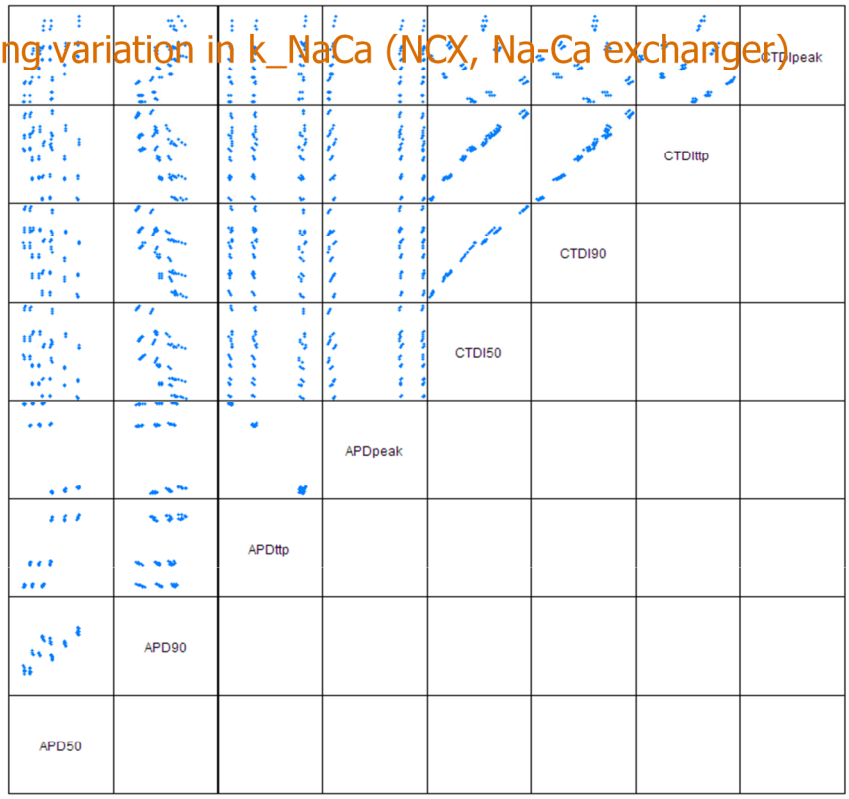
Adding variation in another parameter, we get 3x3 = 9 parameter combinations. Parameter g_Na sets the maximum of the sodium current, whose main role is in the upstroke at the start of the action potential. There isn't much effect on CTD phenotypes, but quite a bit on APD.

Adding variation in vmup_init (SERCA)

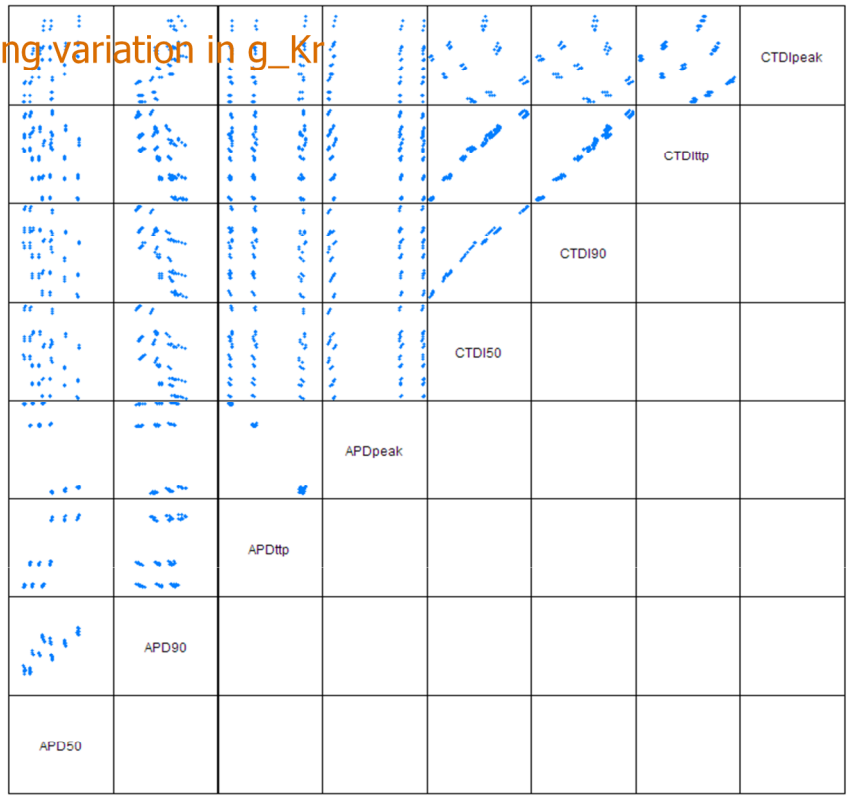


Adding variation in another calcium current: SERCA, inflow to the calcium storage compartments. The 3x3x3=27 parameter scenarios cluster into 3 blobs in the APDttp vs APDpeak panel, 9 blobs for CTDipeak vs CTDittp, and 27 distinct points in the upper left panels.

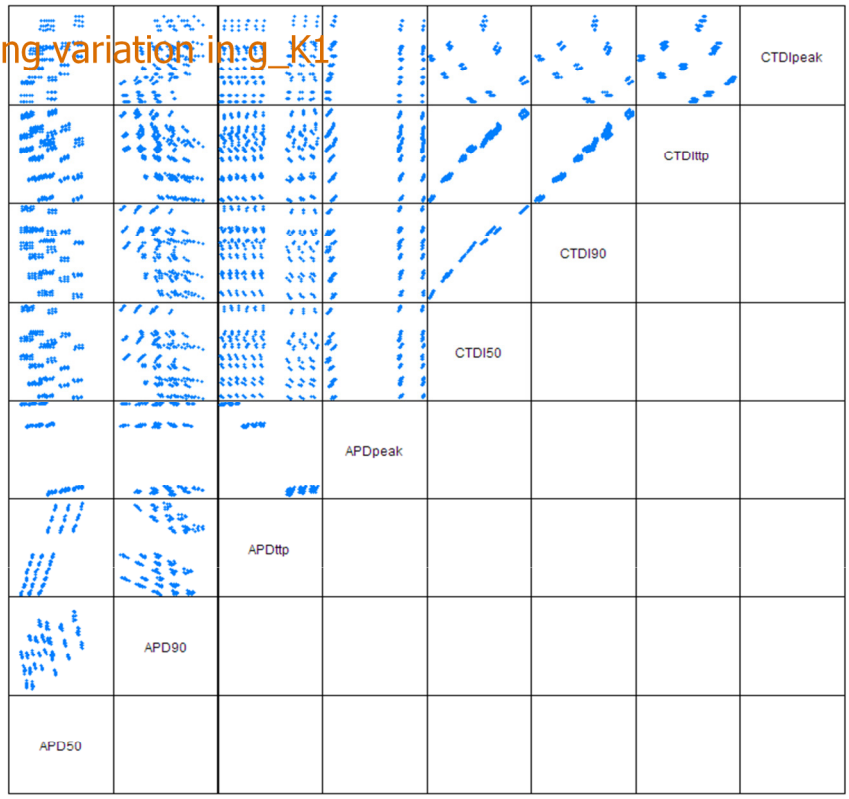
Adding variation in k_NaCa (NCX, Na-Ca exchanger)



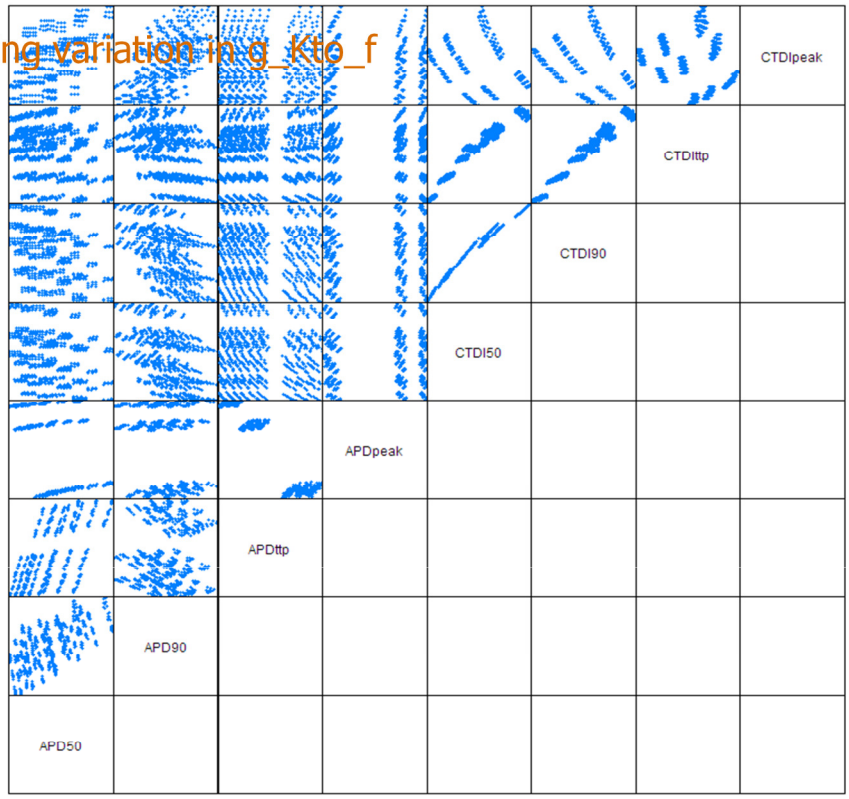
Adding variation in g_Kr

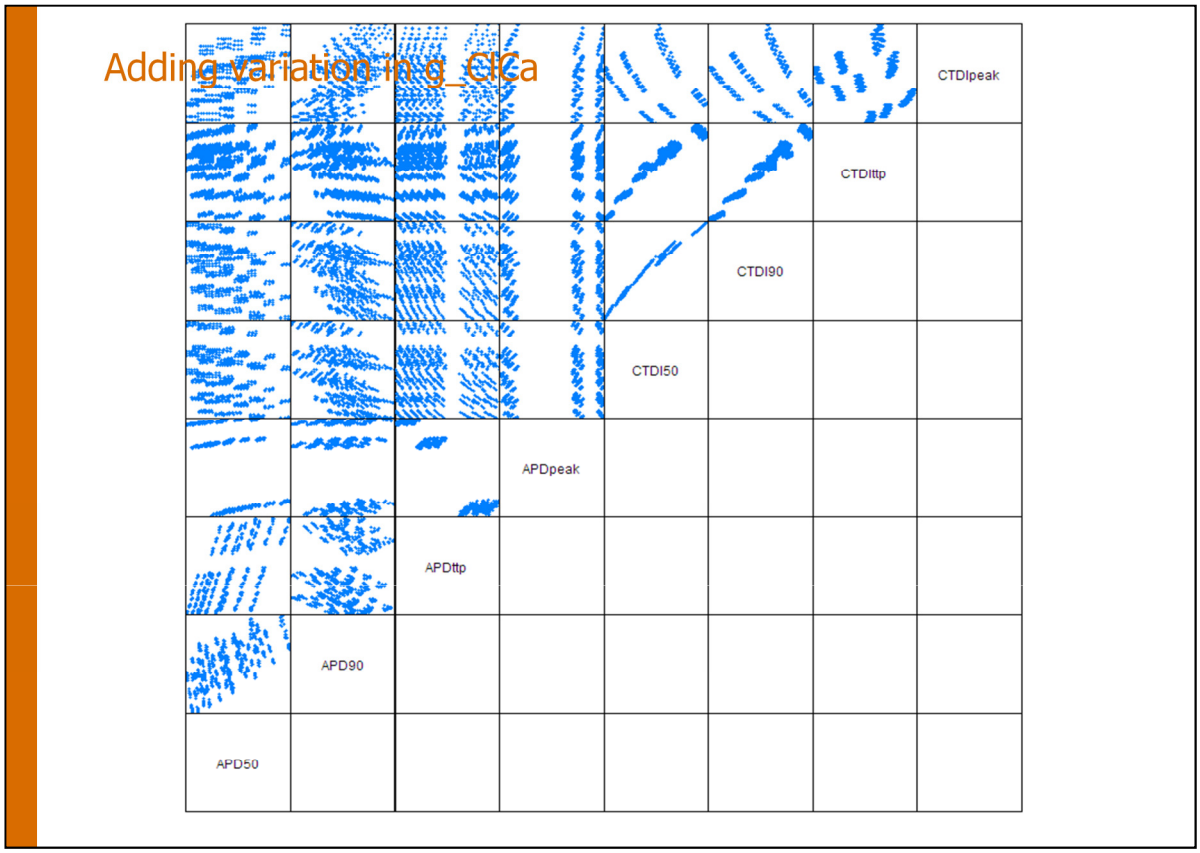


Adding variation in g_K1



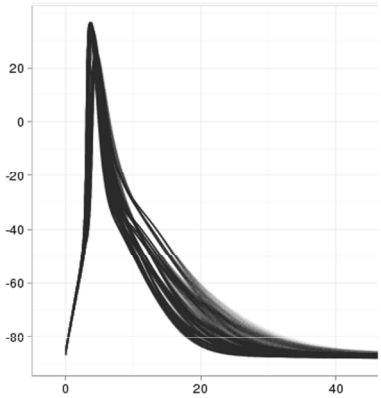
Adding variation in g_Kto_f

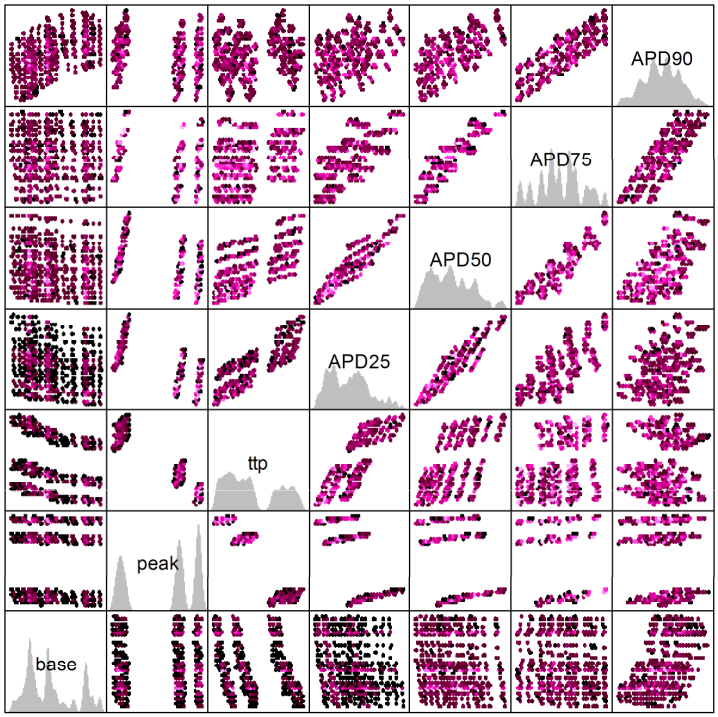




Eventually, 3^8 = 6561 parameter scenarios appear as a fairly smooth distribution for many phenotype pairs, although some phenotypes are clearly determined mostly by one (APDpeak) or two (CTDipeak) of the selected parameters.

Phenotypic correlation



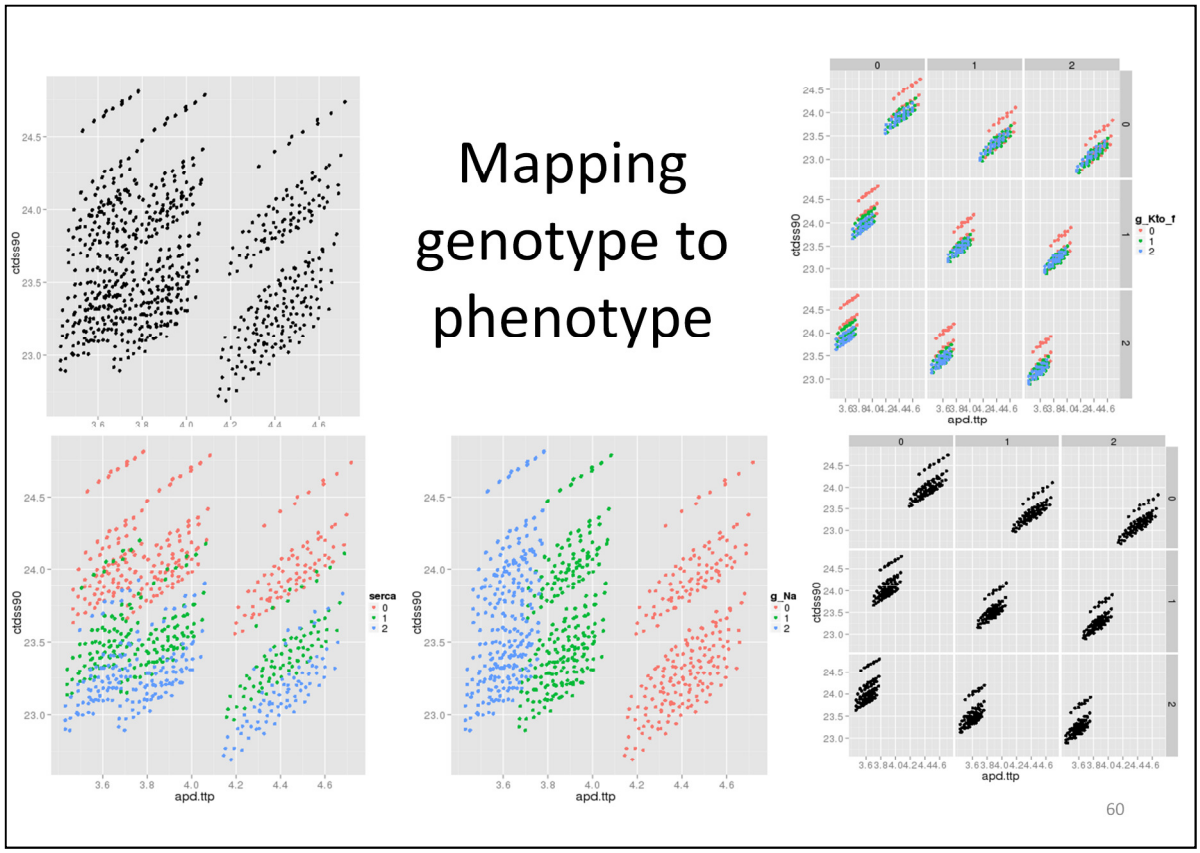


Each panel shows the covariation between two phenotypes, resulting from the simulated genotypic variation in model parameters. Each parameter set is one point in each panel. Here the several thousand points are summarized as densities, progressing from black to magenta.

We see how the added dimension helps account for the marginal distributions [point along diagonal]. For instance, these two peaks [point to ttp] resolve into three groups of points when we consider "peak voltage" and "time to peak" at once.

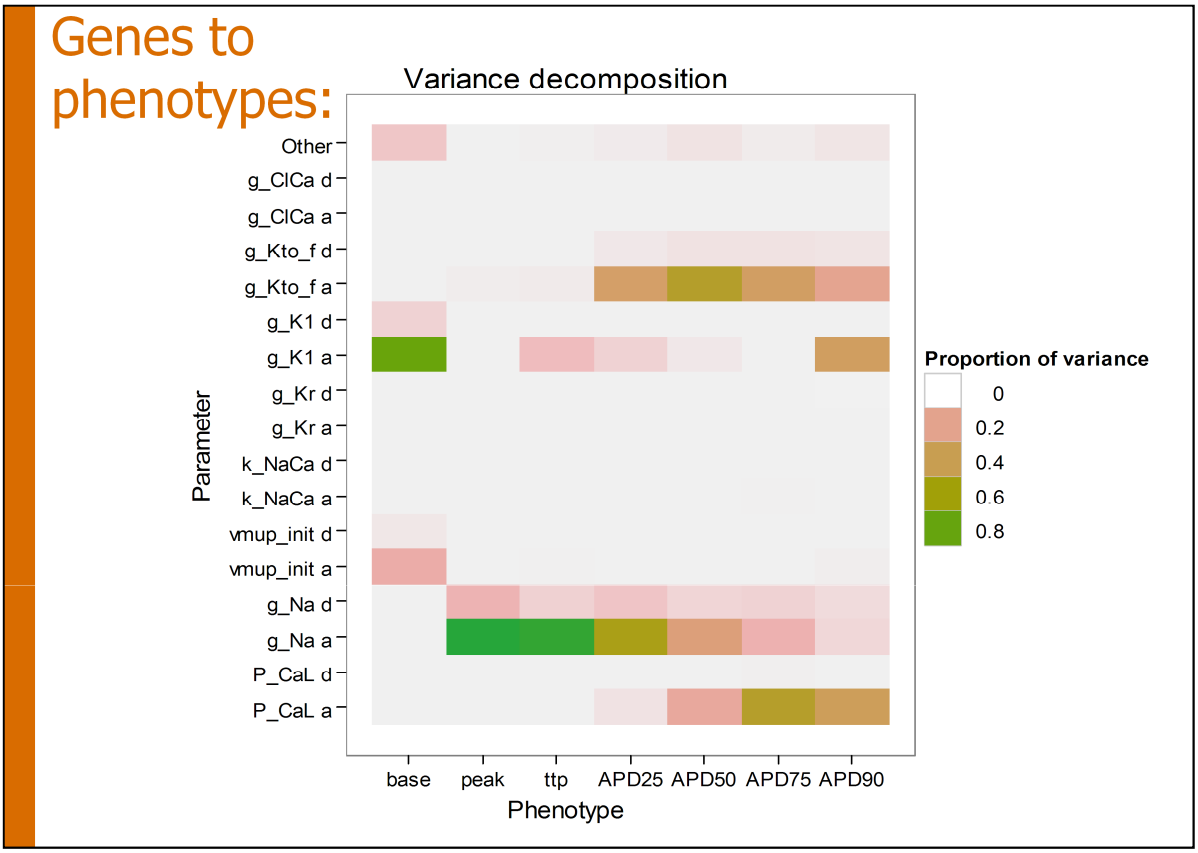
These phenotypes show a strong negative correlation: A fast upstroke gives a tall peak. We also see strong correlation between action potential duration for successive percentages of repolarization [point to first off-diagonal]. However, there is little correlation between the early and late phases [point to ttp vs APD90].

The next slide relates this multivariate phenotype to genes.



This slide zooms in on one pair of phenotypes and shows how the clusters reflect specific "genotypes". (The phenotype on the y axis is calcium transient duration in the "dyadic subspace", which wasn't included on the previous slides. It serves to illustrate the point, though.)

Upper left: uncolored points. Lower left: Points colored by value of SERCA. Lower middle: Points colored by value of g_Na. Lower right: Separate panels for each of the 9 combinations of SERCA x g_Na. Upper right: coloring by g_Kto_f identifies the outlying blob in each panel.



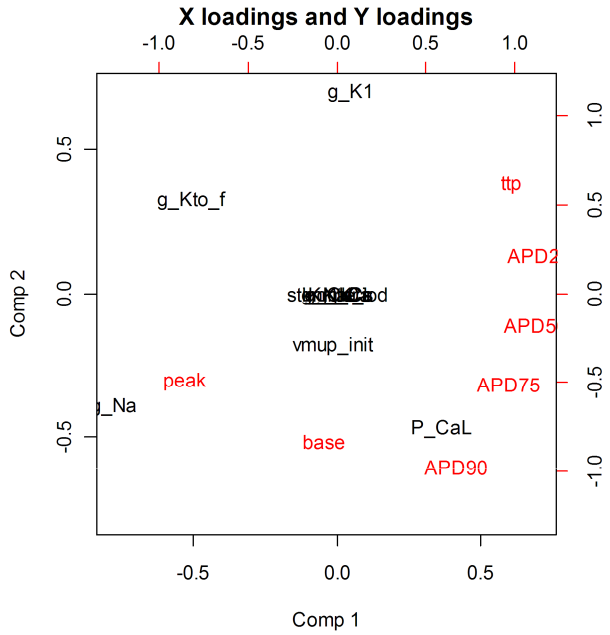
This shows the proportion of phenotypic variance that is accounted for by each genetic parameter. The effects of separate genes generally add up with very little interaction. However, some loci show some dominance.

We recognize "time to peak" as being largely governed by the gene g_Na, with some dominance effect. APD75 turns out to depend on two main genes, plus another with moderate effect. APD90 has three strong effects and two smaller ones.

Conversely, we see that some parameters affect mostly one phenotype, e.g. one phase of the action potential, whereas others affect many. The sodium channel [point to g_Na] governs the initial spike, but that also has consequences for later stages of the action potential. On the other hand, this potassium channel [point to g_Kto_f] affects repolarization only.

Each column in this plot explains variation in a single phenotypic variable, as has been the tradition in quantitative genetics. The next example considers the joint distribution of multiple phenotypes.

Multivariate genotype-phenotype map



Quantitative genetics has until now focused on single-valued phenotypes, one at a time. Dealing with multivariate phenotypes requires novel applications of existing and new statistical tools. For instance, this partial least squares analysis shows that the phenotypes "peak voltage" and "time to peak" are correlated with the parameter g_Na, but in opposite directions. Four "genes" are responsible for most of the phenotypic variation in this virtual experiment.

We recognize the negative correlation between peak amplitude and time to peak. Also, we can see the decreasing correlation between "time to peak" and "action potential duration" as repolarization proceeds [point to APD25-50-75-90].

Genotype-Phenotype: R-squared

	parameter							
phenotype	serca	g_Cab	g_Na	g_Kto_f	g_K1	g_Ks	g_Kur	g_ClCa
apd.peak			91					
apd.ttp		2	88	1	6			
apd25		1	53	32	3		3	
apd50			24	51	1		14	
apd75			2	11			82	
apd90				3	4		88	
ctdss.peak	53	3	9	27			3	
ctdss.ttp	1	1	15	64	4		9	
ctdss25	3	2	8	64	3		14	
ctdss50	8	2	4	56	3		18	
ctdss75	38	5		27	5		16	
ctdss90	64	3	1	9	4		10	
ctdi.peak	73	3	1	15			4	
ctdi.ttp	98							
ctdi25	95			1				
ctdi50	94	1		2				
ctdi75	93	1		3		1		
ctdi90	93	2		3			1	

63

Tredelingsfigurer

Forskjell i additiv effekt på ulike gen og ulike fenotyper

Ørlite eksempel på epistasi

Likevel ser vi igjen struktur i fenotypisk variasjon <=> additivitet

Epistasis? Not much so far…

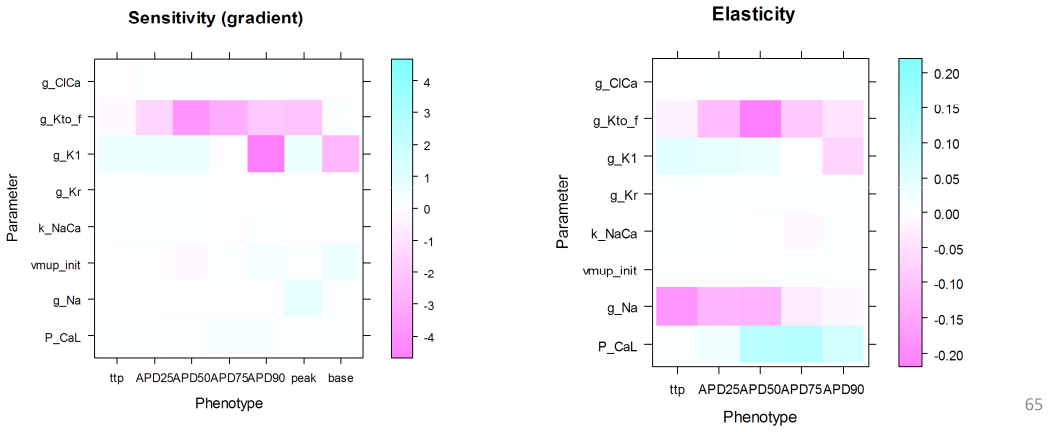
Variance decomposition (percent, ignoring <1%)

	Order 1			Two-gene interactions			
Phenotype	**Total**	A	D	**Total**	AA	AD	DD
APpeak	**100**	92.0	7.9				
APttp	**100**	96.2	3.5				
APD25	**99**	92.7	5.8				
APD50	**95**	90.0	5.0	**5.1**	4.5	0.5	0.0
APD75	**97**	94.2	3.3	**2.5**	2.2	0.3	0.0
APD90	**98**	95.9	2.0	**2.1**	1.9	0.2	0.0
Casspeak	**99**	95.2	4.2				
Cassttp	**97**	94.3	2.9	**2.8**	2.6	0.2	0.0
CassD25	**96**	93.3	3.0	**3.6**	3.3	0.3	0.0
CassD50	**96**	92.1	3.5	**4.4**	4.0	0.4	0.0
CassD75	**95**	90.1	4.9	**5.0**	4.4	0.6	0.0
CassD90	**96**	90.8	5.2	**4.0**	3.5	0.5	0.0
Caipeak	**99**	96.1	3.0	**0.9**	0.9	0.1	0.0
Caittp	**100**	98.2	1.4				
CaiD25	**100**	97.1	2.9				
CaiD50	**100**	96.9	3.0				
CaiD75	**100**	97.6	2.3				
CaiD90	**100**	98.7	1.1				

64

Sensitivity and elasticity

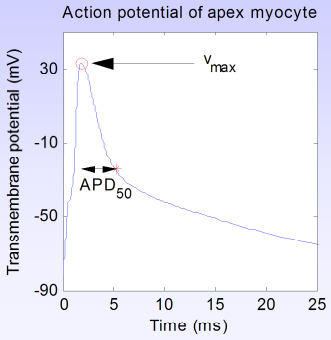
- Sensitivities are df/dxi, where f is a phenotype and xi is a parameter
- Elasticities are (df/f)/(dxi/xi), a dimensionless measure of "% change in phenotype per % change in genotype", assuming the changes are small.
- Phenotypes = columns, genotypic parameters = rows

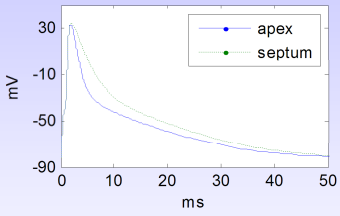


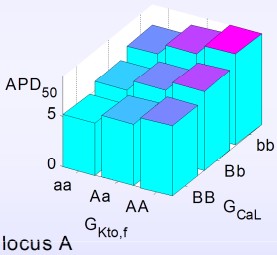
65

Workflow: design, simulate, summarize, analyze

- Link genes to parameters
- Parameter scenarios
- High-dimensional model output
- Summarize phenotype in relevant measures
- Describe and identify patterns: statistical analysis
- Causal explanation: mechanisms in model
- (rinse and repeat)

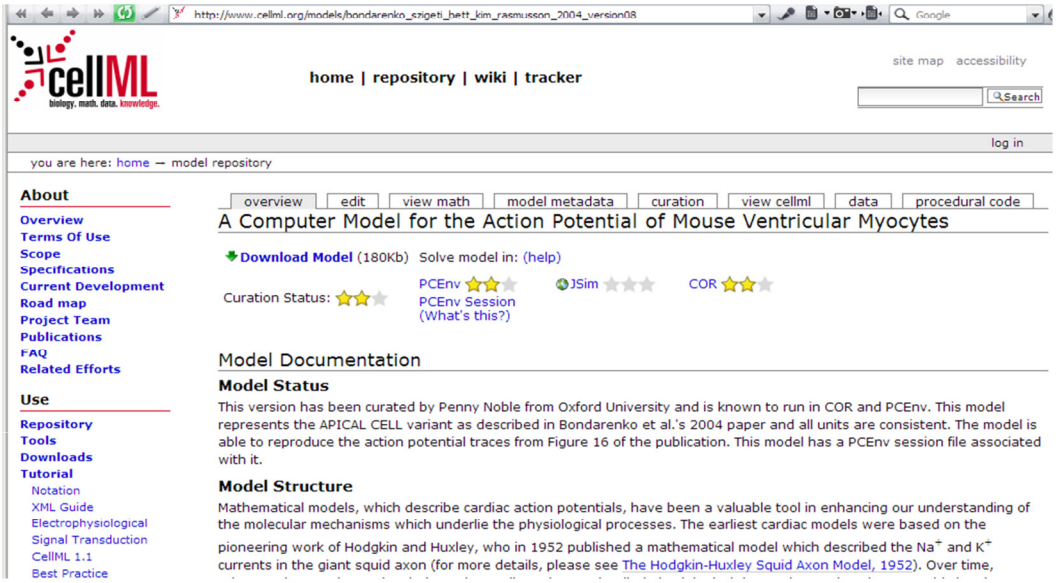






apex	septum		
3.54	6.87	mu	Mean
0.47	0.96	a1	Additive effect of locus A
0.02	0.02	d1	Dominance effect of locus A
-1.15	-0.90	a2	Additive effect of locus B
-0.21	-0.07	d2	Dominance effect of locus B
-0.23	-0.07	iaa	Additive x additive effect of loci A and B
-0.09	0.00	iad	Additive x dominance effect of loci A and B
-0.07	-0.01	ida	Dominance x additive effect of loci A and B
-0.03	-0.04	idd	Dominance x dominance effect of loci A and B

Off-the-shelf, curated physiological models: cellml.org



cGPtoolbox

- Open standards as building blocks
  - HDF5 (used e.g. by NASA) for efficient storage, retrieval, navigation and subsetting of huge data sets
  - CellML model description: curated, consistent, documented (used by the Physiome Project, cellml.org)
- Building on free and open source software
  - Python and Numpy for high-level scripting, streamlined interfaces to HDF5, Sundials (ODE solvers), R (statistical software)
  - NDL, numerical differentiation library: sensitivity and interaction analysis
- cGPtoolbox: In-house development
  - HDF5 conventions for cGP models
  - Virtual genome: Recombination with realistic linkage maps, generating new parameter sets for population studies of cGP models
  - User-friendly, scriptable Python wrappers for CellML, NDL, and Pysundials
  - Multivariate analyses of cGP models, combining existing R packages

Conclusions

- Bridging the gap: relating genetic and physiological observation and modelling, with phenotypic correlation
- Clear causal link between observable physiological things and processes, and phenotypic correlation
- cGP models exhibit similar phenomena and complexity as empirical quantitative-genetic data
- Genetic effects are largely additive. Is life selected for predictable offspring?
- Limits to computing power => need experimental design
  - Must decide where to look; computational resources may be limiting
  - Experimental design

Similarity to empirical work