

# **Phenotypic Variation and Robustness in Complex Genetic**

## **Systems**

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

**Tugce Bilgin Sonay**

aus der Türkei

**Promotionskomitee**

Prof. Dr. Andreas Wagner (Vorsitz)

Prof. Dr. Mark Robinson

Dr. Michael Krützen

**Zürich, 2014**

## Abstract

Phenotypic variation with a genetic basis is the raw material for natural selection and evolution. While it is created by genotypic change, phenotypes are also to some extent robust to genotypic perturbations. Such robustness exists on multiple levels of biological organization. To understand the origins of phenotypic variation with a basis in genetic change, one needs to understand how genotypic changes map to phenotypic changes. In this dissertation I study phenotypic variation and robustness on various levels of biological organizations. I ask how genotypic properties of synthetic metabolisms, such as network size and number of utilizable carbon sources map to metabolic phenotypes, that is, biomass synthesis rates. I describe the trade-offs between these properties quantitatively and show that they can explain most of the variation in synthesis rates of a metabolism. The observations I make are also relevant for synthetic metabolism design, which aims at large-scale, fast, and efficient synthesis of pharmaceuticals, chemical reagents, and biofuels. In a second project, I ask to what extent physicochemical changes in amino acid properties or in protein folding caused by mistranslation affect the codon choice of organisms. I find evidence that selection has increased the incidence of robust codons for ligand-binding amino acids, which suggests that it can affect the robustness of very small units of biological organization. A third project focuses on how tandem repeat instability relates to gene expression divergence in primates. I observe that genes with tandem repeats in gene regulatory regions are associated with high expression divergence. Hence, tandem repeats may contribute substantially to gene expression evolution in primates. Since tandem repeat instability is a hallmark of colorectal tumors, the final project compares phenotypic consequences of tandem repeat instability in gene promoters of tumor and normal tissues. Repeat instability is enhanced in tumors compared to healthy tissues. Those genes with repeat instability are significantly overexpressed. These

findings suggest an important role for tandem repeat instability in the differential gene expression observed for colorectal tumors.

## Zusammenfassung der Dissertation

Phänotypische Variation verursacht durch genetische Variation liefert die Grundlage für die natürliche Selektion und die Evolution. Und obwohl genotypische Veränderungen zu neuen Phänotypen führen können, sind bestehende Phänotypen in gewissem Umfang auch robust gegenüber solchen genotypischen Störungen. Diese Robustheit zeigt sich auf verschiedenen Ebenen der biologischen Organisation eines Organismus. Um den Zusammenhang zwischen phänotypischer und genetischer Variation zu begreifen, muss man insbesondere verstehen, wie genotypische Veränderungen sich auf den Phänotyp auswirken. In dieser Dissertation untersuche ich die phänotypische Variation und Robustheit auf verschiedenen Ebenen der biologischen Organisation eines Organismus. Ich frage danach, wie genotypische Eigenschaften von synthetischen Stoffwechseln, wie zum Beispiel Netzwerkgröße und Anzahl nutzbarer Kohlenstoffquellen, sich auf metabolische Phänotypen auswirken, beispielsweise auf die die Syntheseraten von Biomasse. Ich beschreibe die Kompromisse zwischen diesen sich teilweise widersprechenden genotypischen Eigenschaften, und zeige, dass dadurch der grösste Teil der Variation von Syntheseraten des Stoffwechsels erklärt werden kann. Die Beobachtungen, die ich mache, sind auch für das Design künstlicher Stoffwechsel von Belang, wo es um die schnelle und verlässliche Synthese im grossen Umfang von Pharmazeutika, chemischen Reagenzien, und Biokraftstoffen geht. In einem zweiten Projekt untersuche ich, inwieweit physikalisch-chemische Veränderungen der Eigenschaften von Aminosäuren oder Übersetzungsfehler bei der Proteinfaltung die Auswahl von Codons beeinflussen. Ich finde Hinweise darauf, dass durch natürliche Selektion vermehrt robuste Codons für Aminosäuren verwendet werden, welche Liganden binden. Dies bedeutet, dass natürliche Selektion die Stabilität sehr kleiner Einheiten der biologischen Organisation erhöhen kann. In einem dritten Projekt konzentriere ich mich auf die Instabilität von Tandem-Repeats und untersuche deren Zusammenhang mit den Unterschieden in der

Genexpression bei Primaten. Ich zeige, dass Gene mit Tandem-Repeats in genregulatorischen Regionen mit grossen Unterschieden in der Genexpression verbunden sind. Tandem-Repeats könnten also zur Evolution der Genexpression bei Primaten einen wesentlichen Beitrag leisten. Da die Instabilität von Tandem-Repeats ein Kennzeichen kolorektaler Tumore ist, vergleiche ich im letzten Projekt phänotypische Folgen dieser Instabilität in Gen-Promotoren von Tumorgewebe und von gesunden Gewebe. Die Instabilität von Tandem-Repeats ist in Tumorgewebe erhöht. Jene Gene, deren Instabilität in den Tandem-Repeats deutlich grösser ist werden überexprimiert. Diese Ergebnisse deuten darauf hin, dass die Instabilität von Tandem-Repeats eine wichtige Rolle in der Differenzierung der Genexpression spielt, welche in kolorektalen Tumoren beobachtet werden kann.