

Zusammenfassung

Genomische Variationen in der Modellpflanze *Arabidopsis thaliana* wurden häufig genutzt, um evolutionäre Prozesse in natürlichen Populationen zu verstehen, wobei der Fokus hauptsächlich auf Einzelnukleotid-Polymorphismen (SNPs) lag. Im Gegensatz dazu wurden strukturelle Variationen (SVs) weitgehend ignoriert, obwohl sie das Potential haben, den Phänotyp bedeutend zu beeinflussen. Im ersten Teil meiner Dissertation untersuchte ich unter Verwendung von Illumina-Sequenzierdaten allgemeine Muster struktureller Variationen und Indels in 1301 verschiedenen *A. thaliana*-Varianten aus Marokko, Madeira, Europa, Asien und Nordamerika. Ich überprüfte die identifizierten SVs und verwendete diese in den nachfolgenden Analysen. Ich zeige Belege für eine starke reinigende Selektion von Präsenz-/Abwesenheits-Varianten (PAVs) in Genen, insbesondere für Housekeeping- und Homeobox-Gene. Außerdem fand ich ein Übermaß an PAVs in den der Abwehr zugehörigen Genen (R-Gene, Sekundärmetaboliten) und F-Box-Genen. Dies impliziert das Vorhandensein eines "Kern"-Genoms, das grundlegenden zellulären Prozessen unterliegt, und eines "flexiblen" Genoms, das Gene enthält, die bei räumlich oder zeitlich variierender Selektion wichtig sein könnten. Zusätzlich konnte ich bei fast allen untersuchten Populationen einen Überschuss an PAVs mit mittlerer Frequenz in abwehrreaktionsbezogenen Genen ermitteln, was mit der Geschichte der ausgleichenden Selektion auf diese Klasse von Genen übereinstimmt. Anschließend habe ich Umweltvariablen wie Phänotypen verwendet und gezeigt, dass PAVs zusätzlich zu SNPs dazu beitragen können, mögliche funktionelle genetische Varianten zu identifizieren.

Nach dem Untersuchen des globalen Musters der SVs konzentrierte ich mich im zweiten Teil meiner Arbeit auf die Charakterisierung der strukturellen genetischen Variationen, die sich in der *Arabidopsis thaliana* Variante angehäuft haben, die die Kapverdischen Inseln von Marokko aus kolonisiert hatte. In Übereinstimmung mit nicht veröffentlichten SNP-basierten Analysen fand ich einen drastischen Unterschied in der Anzahl der SVs zwischen Marokko und den Kapverden, wobei die kapverdischen *A. thaliana* Varianten deutlich weniger genetische Variation als die marokkanischen aufwiesen. Mit Hilfe von langen Nanopore-Reads identifizierte ich zwei große SVs, darunter eine 100kb große Inversion auf Chromosom 1 und ein 21kb großes Tandem-Duplikationsereignis (fünfmal wiederholt) auf Chromosom 2. Beide dieser großen SVs überschneiden sich mit Genen, die mit der Blütezeit, der Samenruhe und Stressreaktionen (Licht, Hitze, Trockenheit und Toxine) zusammenhängen. Anschließend habe ich die SVs der *A. thaliana* Variante Cvi-0 dazu verwendet, mögliche funktionelle Varianten zu identifizieren, denen Variationen in

quantitativen Eigenschaften zugrunde liegen. Basierend auf 129 Merkmalen aus 47 Kartierungsstudien und unter Verwendung der Cvi-0 x Ler-0 RIL-Population fand ich potentielle funktionelle Varianten, die sich mit Genen überschneiden, die bei der Samenruhe, Trockenstress, Blütenentwicklung und Blühzeitpunkt beteiligt sind.

Schließlich habe ich erfolgreich zwei Chloroplasten-Haplotypen der *A. thaliana* Variante Cvi-0 assembliert und eine 17 kb Inversion zwischen zwei Haplotypen identifiziert, die sich mit zwei invertierten Regionen und einer kleinen Einzelkopieregion überschneidet.