

INAUGURAL-DISSERTATION

zur Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen Gesamtfakultät
der
Ruprecht-Karls-Universität
Heidelberg

vorgelegt von
Diplom-Biologe Felix Kokocinski
aus Tokyo
2004

Tag der mündlichen Prüfung:

Verarbeitung und funktionelle Analyse
von molekularen Profilen
aus Microarray-Experimenten

Gutachter: PD Dr. Karsten Rippe
Prof. Dr. Peter Lichter

Diese Arbeit wurde am Deutsches Krebsforschungszentrum Heidelberg in der Abteilung Molekulare Genetik unter der wissenschaftlichen Leitung von PD Dr. Karsten Rippe und Prof. Dr. Peter Lichter durchgeführt.

Ich möchte mich an dieser Stelle besonders bei Prof. Dr. Peter Lichter bedanken. Er ermöglichte die vorliegende Arbeit und ist in vielerlei Hinsicht für mich eine vorbildliche Führungspersönlichkeit.

Mein herzlichster Dank gilt Dr. Anja Kolb-Kokocinski, die mich in der oft knappen Zeit sehr unterstützt hat, obwohl sich auch Ihre eigene Arbeit stapelte.

Eine großartige Unterstützung war mir Dr. Gunnar Wrobel, ohne dessen Genialität und Motivation kaum eines der hier vorgestellten Projekte zur heutigen Reife gelangt wäre.

Ein Gruß und Dank an meine Freiburger und Heidelberger Studienfreunde, die wichtige Wegbegleiter waren und dies auch immer bleiben werden.

Durch die wertvollen Kollegen und Freunde der Arbeitsgruppe *Molekulare Genetik* war das Arbeitsumfeld für mich immer sehr angenehm und motivierend, auch hierfür ein großer Dank, besonders an Grischa Tödt und Lars Hummerich.

Ein weiterer Dank geht an die fleißigen Korrekturleser.

Inhalt	Seite
1. Einleitung	1
1.1. DNA-Microarray-Experimente	1
1.2. Datenverarbeitung in Hochdurchsatz-Experimenten	5
2. Problemstellungen und Relevanz	7
2.1. Entwicklung eines Managementsystems für Klon-Bibliotheken	7
2.2. Kontrolle und Archivierung des Prozessablaufs zur Herstellung von Microarrays	8
2.3. Generierung von Primersequenzen für <i>Real-Time PCR</i> Experimente	9
2.4. Computergestützte Analyse von Microarray-Ergebnissen	10
2.5. Untersuchung der Pathomechanismen von <i>Non-Melanom</i> Hautkrebs	11
3. Material und Methoden	13
3.1. Das Datenbanksystem <i>CloneBase</i>	13
3.2. Das Prozessdatensystem <i>QuickLIMS</i>	14
3.3. Das Primergenerierungs-Werkzeug <i>AutoPrime</i>	15
3.4. Das Analyseprogramm <i>FACT</i>	16
3.5. Microarray-Experimente	17
4. Ergebnisse	19
4.1. Die Klon-Datenbank <i>CloneBase</i>	19
4.1.1. <i>CloneY</i> - cDNA Klone und Oligomere für Expressions-Microarrays	19
4.1.2. <i>CloneX</i> - genomische Klone für Matrix-CGH-Experimente	24
4.1.3. Datenaktualisierung mit dem <i>DB-Updater</i>	25
4.1.4. Datenaktualisierung mit automatisierten Skripten	30
4.1.5. Internet-Abfrageseiten für den Benutzer (<i>Web-Interface</i>)	32
4.2. Die Prozess-Datenbank <i>QuickLIMS</i>	37
4.2.1. Grundsätzlicher Aufbau	37
4.2.2. Spezifische Funktionen	39
4.2.3. Interaktion mit dem Pipetier-Roboter und der Klondatenbank	41
4.3. Das <i>AutoPrime</i> Programm zur automatisierten Primergenerierung	42
4.4. Funktionelle Analyse von Experimenten (<i>FACT</i>)	47
4.4.1. Grundsätzlicher Aufbau	47
4.4.2. Die Datenbank-Basis	48
4.4.3. Flexible Integration heterogener Datenquellen	51
4.4.4. Analyse mit Annotationsdaten	54

4.4.5. Ausgabe von Ergebnissen	57
4.4.6. Internet-Oberfläche	59
4.5. Untersuchung der Entstehung und Progression von <i>Non-Melanom</i> Hautkrebs	61
4.5.1. Durchführung der Experimente	61
4.5.2. Datenanalyse	61
4.6. Weitere Anwendungen im Bereich der Krebsforschung	64
5. Diskussion	68
5.1. Das Datenbanksystem <i>CloneBase</i>	68
5.2. Das Prozessdatensystem <i>QuickLIMS</i>	71
5.3. Das Primergenerierungs-Werkzeug <i>AutoPrime</i>	73
5.4. Das Analyseprogramm <i>FACT</i>	75
5.5. Untersuchung der Pathomechanismen von <i>Non-Melanom</i> Hautkrebs	78
5.6. Das Gesamtsystem	80
6. Literaturverzeichnis	82
7. Anhang	90
7.1. Erläuterung von Abkürzungen und Fachtermini	90
7.2. Eigene Publikationen	92
7.2.1. Publikationen in Fachzeitschriften	92
7.2.2. Beiträge zu internationalen Konferenzen	93
8. Anhang auf CD-ROM (bzw. unter http://www.dkfz.de/molgen/div/kokocinski):	94
Programmcode und Datenbanken	
8.1. <i>CloneBase</i> (mySQL-Datenbank, Java Programm <i>DB-Updater</i> , Perl Update-Skripte)	
8.2. <i>QuickLIMS</i> (Access Datenbank mit VBA-Programmcode)	
8.3. <i>AutoPrime</i> (Perl Hauptprogramm, <i>Web-Interface</i> und Zusatz-Dateien)	
8.4. <i>FACT</i> (mySQL Datenbank, Perl API-Code und <i>Web-Interface</i>)	

Verzeichnis der Abbildungen

Abb. 1: DNA Microarrays	2
Abb. 2: Mausmodell zur chemisch-induzierten Karzinogenese	12
Abb. 3: Schema der Datenbank <i>CloneBase</i>	21
Abb. 4: Ansicht der Programm-Oberfläche vom <i>DB-Updater</i>	27
Abb. 5: Die unterschiedlichen Datenquellen der <i>CloneY</i> -Tabelle	28
Abb. 6: Aktualisierungsvorgang der <i>CloneBase</i> mittels automatisierter Perl-Skripte	31
Abb. 7: Eingesetzte Client-Server-Architektur und Technologie	32
Abb. 8: <i>CloneBase</i> Web-Oberfläche	34
Abb. 9: Internet-Abfrageseiten des <i>CloneBase</i> -Projekts	36
Abb. 10: Platten-orientierter Aufbau von <i>QuickLIMS</i>	37
Abb. 11: <i>QuickLIMS</i> -Hauptmaske für Benutzer-Interaktionen	40
Abb. 12: Funktionsweise von <i>AutoPrime</i>	43
Abb. 13: Aufbau des <i>AutoPrime</i> -Programms	44
Abb. 14: Internet-Oberfläche von <i>AutoPrime</i>	46
Abb. 15: Verwendung von heterogenen Datenquellen mit <i>FACT</i>	47
Abb. 16: Aufbau des <i>FACT</i> -Systems	48
Abb. 17: Datenbank-Schema von <i>FACT</i>	50
Abb. 18: Beispiel für Annotations-Stufen	51
Abb. 19: Beispiele für von <i>FACT</i> -Modulen produzierte Ergebnis-Graphiken	58
Abb. 20: Webinterface von <i>FACT</i>	60
Abb. 21: <i>FACT</i> -Analysen zur Untersuchung des <i>Non-Melanom</i> Hautkrebses	63
Abb. 22: <i>FACT</i> -Analysen zur Untersuchung des <i>Non-Melanom</i> Hautkrebses	64
Abb. 23: Karte der 2 MB <i>EDC</i> -Region in der chromosomalen Bande 1q21	79

Verzeichnis der Tabellen:

Tab. 1: Durchführung eines Microarray-Experiments und dazugehörige Fragen an die Datenverarbeitung	6
Tab. 2: Zusammensetzung der Klonsammlung in der Tabelle <i>CloneY</i>	20
Tab. 3: Beschreibung der Felder der <i>CloneY</i> -Tabelle	22
Tab. 4: Beschreibung der Felder der <i>Genes</i> -Tabelle	23
Tab. 5: Beschreibung der Felder der <i>CloneX</i> -Tabelle	24
Tab. 6: Zusammensetzung der Klonsammlung der Tabelle <i>CloneX</i>	25
Tab. 7: Zur Verfügung stehende Datenquellen für die Annotation in <i>FACT</i>	52
Tab. 8: Analysefunktionen, die in <i>FACT</i> genutzt werden können	55
Tab. 9: Anwendung der unterschiedlichen Systeme in konkreten Forschungsprojekten	65