

Inhaltsverzeichnis

1	Einleitung	1
1.1	Bioinformatik	1
1.2	Systembiologie	3
1.3	Der Sonderforschungsbereich 578: Vom Gen zum Produkt	4
1.3.1	Das Teilprojekt B4: Systembiologie der Produkt- und Pelletbildung durch <i>Aspergillus niger</i>	5
1.4	Die Schimmelpilzgattung <i>Aspergillus</i>	6
1.4.1	<i>Aspergillus niger</i>	7
1.4.2	Weitere <i>Aspergillus</i> -Arten	7
1.5	Heterologe Proteinexpression	8
1.5.1	Kodonnutzung und heterologe Proteinproduktion	10
1.5.2	Kodonadaptierung	11
1.6	Aufgabenstellung	13
2	Material und Methoden	15
2.1	Annotation von Genomen	15
2.1.1	<u>G</u> ene <u>L</u> ocator and <u>I</u> nterpolated <u>M</u> arkov <u>M</u> odel <u>E</u> R (GLIMMER)	15
2.1.2	<u>B</u> asic <u>L</u> ocal <u>A</u> lignment <u>S</u> earch <u>T</u> ool (BLAST)	16
2.1.3	<u>M</u> etabolic <u>S</u> earch and <u>R</u> econstruction <u>K</u> it (metaSHARK)	17
2.1.4	tRNAscan-SE	18
2.2	Datenquellen der <i>Aspergillus</i> -Datenbank	19
2.2.1	Quellen der Genomdaten verschiedener <i>Aspergillus</i> -Arten	19
2.2.1.1	Integrated Genomics	19
2.2.1.2	<u>J</u> oint <u>G</u> enome <u>I</u> nstitute (JGI)	20
2.2.1.3	DSM	20
2.2.1.4	Broad Institute	21
2.2.2	Weitere Quellen für die Erstellung der <i>Aspergillus</i> -Datenbank	21
2.2.2.1	Die „ <u>K</u> yoto <u>E</u> ncyclopedia of <u>G</u> enes and <u>G</u> enomes (KEGG)“	22

2.2.2.2	Swiss-Prot und TrEMBL	22
2.3	Optimierung heterologer Proteinexpression durch Kodonadaptierung	23
2.3.1	<u>C</u> odon <u>A</u> daptation <u>I</u> ndex (CAI)	23
2.3.2	Algorithmus zur iterativen Berechnung des CAI	24
2.3.3	Algorithmus zur Vorhersage von rho-unabhängigen Transkriptions-Terminatoren	26
2.4	Datenquellen für die Kodonanpassung	27
2.4.1	Die Genome-Reviews-Datenbank	28
2.4.2	Weitere Datenquellen	28
2.4.3	Translationstabellen vom „National Center of Biotechnology Information (NCBI)“	30
2.4.4	The <u>R</u> estriction <u>E</u> nzyme <u>D</u> atabase (REBASE)	30
2.5	Verwendete Techniken aus der Informatik	31
2.5.1	Das Datenbankmanagementsystem PostgreSQL	31
2.5.2	Java	33
2.5.3	Das „Java-Package“ JFreeChart	33
2.5.4	Das Java-Datenbankmanagementsystem HSQLDB	34
2.5.5	<u>J</u> ava <u>S</u> erver <u>P</u> ages (JSP)	35
2.5.6	Apache Tomcat	35
2.5.7	<u>S</u> imple <u>O</u> bject <u>A</u> ccess <u>P</u> rotocol (SOAP)	35
3	Ergebnisse und Diskussion	39
3.1	Die <i>Aspergillus</i> -Datenbank „ANigerDB“	39
3.1.1	Formale Eigenschaften der <i>Aspergillus</i> -Datenbank	39
3.1.2	Annotation der <i>Aspergillus niger</i> Stämme ATCC 1015 und NRRL3 mit verschiedenen Methoden	42
3.1.2.1	Annotation mit Hilfe von BLAST und Glimmer-HMM	42
3.1.2.2	Annotation mit Hilfe von metaSHARK	44
3.1.2.3	Integration der Daten aus der KEGG-Datenbank	44
3.1.3	Vergleich der Annotationen der <i>Aspergillus niger</i> -Stämme	44
3.1.4	Ein kurzer Vergleich der unterschiedlichen <i>Aspergillus</i> -Arten	48
3.1.5	Experimentelle Daten in der Datenbank	50

3.1.6	Webinterface der <i>Aspergillus</i> -Datenbank	51
3.1.7	Kurze Zusammenfassung der <i>Aspergillus</i> -Datenbank „ANigerDB“	55
3.2	<u>J</u> ava <u>C</u> odon <u>A</u> daptation <u>T</u> ool (JCat)	56
3.2.1	Anpassung der „codon usage“	56
3.2.2	Weitere Optionen bei der Anpassung der „codon usage“	57
3.2.2.1	Der genetische Code der Eingabesequenz	58
3.2.2.2	Vermeidung Rho-unabhängiger Transkriptionsterminatoren	58
3.2.2.3	Vermeidung prokaryotischer Ribosomenbindestellen (Shine-Dalgarno-Sequenz)	60
3.2.2.4	Vermeidung spezieller Restriktionsenzymbindestellen	61
3.2.2.5	Unvollständige Anpassung der „codon usage“ der eingegebenen Sequenz	61
3.2.3	Berechnung von CAIs aus einer Datei im FASTA-Format	62
3.2.4	Das Datenbankmodell von JCat	62
3.2.5	Anwendungsoberflächen von JCat	63
3.2.5.1	Der JCat-Webserver	64
3.2.5.2	JCat als eigenständiges Programm	65
3.2.6	Anwendungsbeispiele für JCat aus der Literatur	65
3.2.7	Anpassung des <i>Escherichia coli</i> Arsenat Reduktase Gens <i>arsC</i> an die „codon usage“ von <i>Bacillus megaterium</i>	65
3.2.8	Kurze Zusammenfassung des Programms „JCat“	66
4	Ausblick	69
5	Zusammenfassung	71
6	Anhang	73
7	Abkürzungsverzeichnis	79
8	Literaturverzeichnis	81