



UPPSALA  
UNIVERSITET

## Molecular Biotechnology Programme

Uppsala University School of Engineering

<b>UPTEC X 10 021</b>		<b>Date of issue 2010-11</b>
Author <p style="text-align: center;"><b>Lennie Fredriksson</b></p>		
Title (English) <p style="text-align: center;"><b>Clustering of DNA sequence reads from repeat regions using defined nucleotide positions (DNPs)</b></p>		
Title (Swedish)		
Abstract <p>Sequencing genomes with a high frequency of repeat regions is a difficult task. The aim of the project was to develop an algorithm to speed up the sequencing process of highly repetitive genome. By using specific differences between the repeats called defined nucleotide positions (DNPs), cluster DNA sequence reads into contigs. The strategy used in the development of the algorithm resulted in a quite complex algorithm. Test runs of the algorithm showed that there is still work to be done to get a desirable result.</p>		
Keywords <p>DNP, repeat regions, SolidClusters, algorithm</p>		
Supervisors <p style="text-align: center;"><b>Erik Arner</b> Karolinska Institutet, Center for Genomics and Bioinformatics</p>		
Scientific reviewer <p style="text-align: center;"><b>Siv Andersson</b> Uppsala Universitet, Evolutionsbiologiskt Centrum</p>		
Project name	Sponsors	
Language <p style="text-align: center;"><b>English</b></p>	Security	
<b>ISSN 1401-2138</b>	Classification	
Supplementary bibliographical information	Pages <p style="text-align: center;"><b>45</b></p>	
<b>Biology Education Centre</b> Box 592 S-75124 Uppsala	Biomedical Center Tel +46 (0)18 4710000	Husargatan 3 Uppsala Fax +46 (0)18 471 4687