

Pyrosequencing™ Assay Design Software

User Manual



Legal

Warranty and Liability


Biotage AB warrants that the product supplied has been thoroughly tested to ensure that it meets its published specifications. The warranty is only valid if the product has been installed and used according to the instructions provided by Biotage AB.

Biotage AB makes no warranties, expressed or implied, including without limitation the implied warranties of merchantability and fitness for a particular purpose regarding the product. Biotage AB does not warrant, guarantee or make any representations regarding the use or the results of the use of the product in terms of its correctness, accuracy, reliability, currentness or otherwise. The user assumes the entire risk, as to the results and performance of the product. Since the exclusion of implied warranties is not permitted by some jurisdictions, the above exclusion may not necessarily apply.

Biotage AB shall in no event be liable for any direct, indirect, special or consequential damages including, without limitation, damages for loss of business income, business profits, business interruption, loss of business information and the like arising out of the use or inability to use the product.

Since the exclusion of implied warranties is not permitted by some jurisdictions, the above exclusion may not necessarily apply.

Trademarks and patents owned by Biotage AB

Pyrosequencing, PSQ, Pyrogram and  are trademarks owned by Biotage AB. Pyrosequencing technology is covered by patents including patents US4863849, US610891, US6258568, EP0932700, and EP0946752, and patent applications owned by Biotage AB.

In view of the risk of trademark degeneration, authors intending to use the trademarked designations are respectfully requested to acknowledge the trademark status of the products at least once in each article.

Other patents and trademarks

The PCR process is covered by several patents owned by Roche Molecular Systems and F.Hoffman-La Roche Ltd.

Intel and Pentium are registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Microsoft and Windows are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

Sepharose is a trademark of Amersham Biosciences Limited.

Adobe and Adobe Acrobat are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.

All other trademarks are the property of their respective owners.

Important user information

The Pyrosequencing™ Assay Design Software and all associated products from Biotage AB are for research purposes only. Not for use in diagnostics procedures for clinical purposes. For in vitro use only.

Biotage AB reserves the right to make changes in the information contained herein without prior notice.

Pyrosequencing Assay Design Software, User Manual version 1.0.6 AA.

© Copyright 2004 Biotage AB

All rights reserved. No part of this manual may be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without the expressed written permission of Biotage AB.

Biotage AB
Kungsgatan 76
SE-753 18 Uppsala
SWEDEN
Phone: +46 18 56 59 00
Fax: +46 18 59 19 22
E-mail: info@biotage.com
Web: <http://www.biotage.com>

Contents

1	Introduction.....	5
1.1	The User Manual	5
1.2	The Quick Guide.....	5
2	Software setup.....	6
2.1	System requirements.....	6
2.2	Installing the software	6
2.3	Starting Pyrosequencing Assay Design Software	9
2.4	The Assay Design Software start screen	10
3	Assay design settings	11
3.1	Change default assay design settings	11
3.2	Change assay design settings for one assay	11
3.3	Description of settings and buttons	12
3.3.1	<i>PCR Primer settings.....</i>	<i>12</i>
3.3.2	<i>Sequencing primer settings.....</i>	<i>12</i>
3.3.3	<i>Primer set settings.....</i>	<i>13</i>
3.3.4	<i>Buttons.....</i>	<i>13</i>
3.4	Selected settings will affect primer set scoring.....	13
3.4.1	<i>PCR Primer score effects.....</i>	<i>13</i>
3.4.2	<i>Sequencing primer score effects.....</i>	<i>14</i>
4	Performing an assay design	15
4.1	Introduction	15
4.2	Create an assay setup.....	16
4.2.1	<i>Step 1 - Choose assay type and enter a description for the assay</i>	<i>16</i>
4.2.2	<i>Step 2 - Enter the DNA sequence</i>	<i>16</i>
4.2.3	<i>Step 3- Name Polymorphisms and unknown sequences</i>	<i>25</i>
4.2.4	<i>Step 4- Optional: Set the target region</i>	<i>25</i>
4.2.5	<i>Step 5 - Optional: Redefine PCR primer regions.....</i>	<i>27</i>
4.3	Run an automatic design	27
4.3.1	<i>Generate new primer sets.....</i>	<i>27</i>
4.3.2	<i>Analyze previously designed primers</i>	<i>28</i>
4.4	View results.....	30
4.4.1	<i>View results - overview.....</i>	<i>30</i>
4.4.2	<i>Scoring and quality</i>	<i>30</i>
4.4.3	<i>Sort the primer set list</i>	<i>31</i>
4.4.4	<i>View information on different primer sets</i>	<i>32</i>
4.4.5	<i>Optional: Select a different primer set as final.....</i>	<i>33</i>
4.5	Adjust an assay	33
4.5.1	<i>Change the assay design settings.....</i>	<i>33</i>
4.5.2	<i>Edit the PCR and sequencing primers.....</i>	<i>34</i>
4.5.3	<i>Select one or more primers from a primer set and re-analyze the assay</i>	<i>35</i>
4.6	Save an assay	35
4.7	View an assay report	36
4.7.1	<i>View a report</i>	<i>36</i>
4.7.2	<i>Print a report</i>	<i>37</i>
4.7.3	<i>Save a report.....</i>	<i>37</i>
5	Performing batch assay design	38
5.1	Introduction	38
5.2	Run automatic batch assay design	39
6	Importing an assay file into PSQ system software.....	40
7	Guidelines for PCR and sequencing primer design	41
7.1	Introduction	41
7.2	Analysis steps performed by the software	41
7.2.1	<i>Introduction.....</i>	<i>41</i>
7.2.2	<i>Warning messages.....</i>	<i>43</i>
7.2.3	<i>PCR primer analyses</i>	<i>45</i>

7.2.4	PCR primer pair analyses	46
7.2.5	Sequencing primer analyses	47
7.2.6	Primer set analyses	48
7.3	Melting temperature	49
7.3.1	Methods for calculating the melting temperature (T_m)	49
7.4	Guidelines for PCR primer design	49
7.4.1	General guidelines to apply when designing PCR primers	50
7.5	Guidelines for sequencing primer design	51
7.5.1	General guidelines for sequencing primer design	51
7.5.2	Sequencing primer design for genotyping and allele quantification	52
7.5.3	Sequencing primer design for SQA	53
8	Guidelines for PCR setup and optimization	54
8.1	Guidelines for PCR setup and optimization for Pyrosequencing analysis	54
8.2	PCR setup and optimization - Specific for allele quantification	56
8.3	PCR protocol example	57
8.3.1	Optimization of the PCR protocol and conditions	57
9	Hints & Tips	58
9.1	Tips for succeeding with difficult assays	58
9.2	Tips for avoiding PCR cross-contamination	59
9.3	Tips for assay controls	59
9.4	Tips for multiplex assay design	60
9.5	Tips for universal biotinylated PCR primers	60
9.6	Tips for using the Gibbs free energy (ΔG) graph	61
9.7	Tips for using the melting temperature (T_m) graph	62
9.8	Tips for analyzing InDels in homopolymeric stretches	62
9.9	Analyzing short tandem repeats (STRs)	63
9.10	Troubleshooting guide	64
10	Appendix A. Methodological background	66
10.1	Sample preparation	66
10.2	Pyrosequencing systems	67
10.2.1	Introduction	67
10.2.2	Genotyping and mutation analysis	67
10.2.3	Allele quantification (AQ)	67
10.2.4	Sequence analysis (SQA)	68
10.3	Definitions	68
10.3.1	Alleles	68
10.3.2	Single nucleotide polymorphisms (SNPs)	69
10.3.3	Insertions and deletions (InDels)	69
10.3.4	Short tandem repeats (STRs)	70
10.3.5	Sequence database files	70
10.3.6	Sequence to analyze (genotyping and allele quantification)	71
10.3.7	Dispensation order	71
10.3.8	Reference peaks and quality control window	72
10.3.9	Simplex and multiplex assays	72
10.3.10	Mispriming	73
10.3.11	Secondary structures	73
11	Appendix B. Assay types	75
11.1	Genotyping and allele quantification (AQ)	75
11.1.1	Introduction	75
11.1.2	Polymorphisms for which assays can be designed	75
11.1.3	Entering polymorphisms in Assay Design Software	77
11.2	Sequence analysis (SQA)	78
11.2.1	Introduction	78
11.2.2	Entering sequences for SQA into Assay Design software	78
12	Glossary	79
13	Index	85

1 Introduction

The Pyrosequencing™ Assay Design Software is a tool for designing PCR and sequencing primers for Pyrosequencing assays. Applications supported by the software are genotyping, mutation analysis, allele quantification (AQ), and sequence analysis (SQA).

The software generates primer sets (each primer set consists of a PCR primer pair and a sequencing primer) that fulfill the specific requirements for Pyrosequencing analysis.

1.1 The User Manual

The User Manual contains user instructions for the software, guidelines for PCR and sequencing primer design, guidelines for PCR setup and optimization, and hints & tips for designing assays. It also contains instructions for installing the software and computer requirements.

The appendix of the User Manual contains information on the Principle of Pyrosequencing, assay types and general information on sample preparation, Pyrosequencing systems and basic definitions.

1.2 The Quick Guide

The Quick Guide contains short, concise instructions for performing a typical assay design. The Quick Guide is available as a separate document on the installation CD. When the installation CD is inserted into the CD-drive, an installation wizard will automatically start. The wizard presents three different choices: install Assay Design software, view the User Manual, or view the Quick Guide. If the wizard does not start automatically, the CD-drive can be opened in Windows Explorer and the wizard started by double-clicking on the autorun.exe file.



2 Software setup

These installation guidelines describe how to install Pyrosequencing Assay Design Software on a computer fulfilling the system requirements listed below. The software usage is restricted by a license key. The license key information, which can be found on a slip in the installation CD folder, is needed when the software is started for the first time.

2.1 System requirements

Computer

The computer used to run Assay Design Software should be a PC with the following preferred specifications.

Processor	2.6 GHz
RAM	1 GB
Hard drive	100 MB free space
Monitor	1024 x 768 resolution, Medium color quality (16bit)
Operating system	Microsoft® Windows® 2000 or Microsoft® Windows® XP, English versions only

Printer

All printers supported by Windows 2000 or XP are suitable.

Backup of data

Good data management requires that data backups be made on a regular basis.

Note: Biotage AB is not responsible for the User's backup routines.

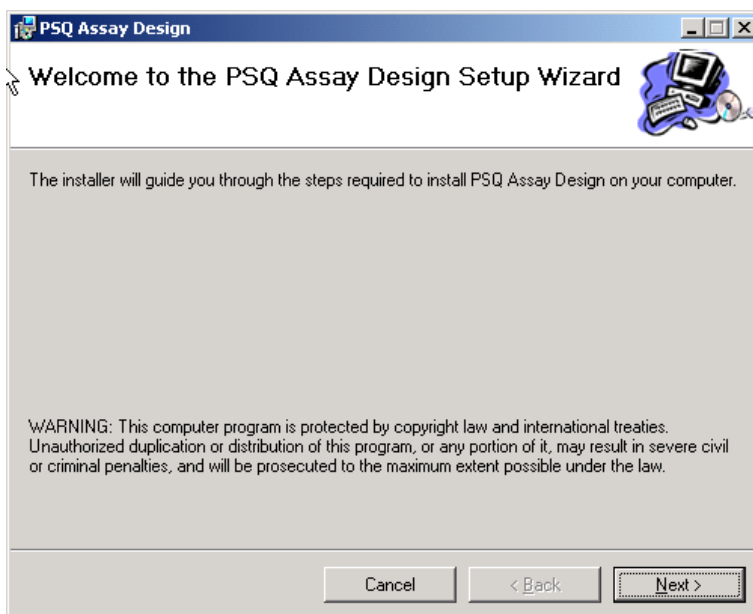
2.2 Installing the software

1. Before starting the installation of Pyrosequencing Assay Design Software, confirm that you have administrator's rights on the computer.
2. Insert the Pyrosequencing Assay Design Software CD into the CD-drive of the computer.

3. Follow the Assay Design Setup wizard, which automatically starts. If the wizard does not start, either open the CD-drive in Windows Explorer and double-click the autorun.exe file, or choose **Run** in the Windows **Start** menu, specify the path to the CD-drive and the file autorun.exe, e.g. D:\autorun.exe



4. Pyrosequencing Assay Design Software requires Microsoft® .NET Framework version 1.1 to be installed and running on the computer. The installation wizard will automatically detect the presence of .NET Framework or will, if necessary, install the program. If the installation of .NET Framework fails, the installation will stop with an error message. Confirm that you are logged in with administrator rights and restart the installation.
5. After installation of .NET Framework, installation continues with Pyrosequencing Assay Design Software. Click **Next** to proceed with the installation.

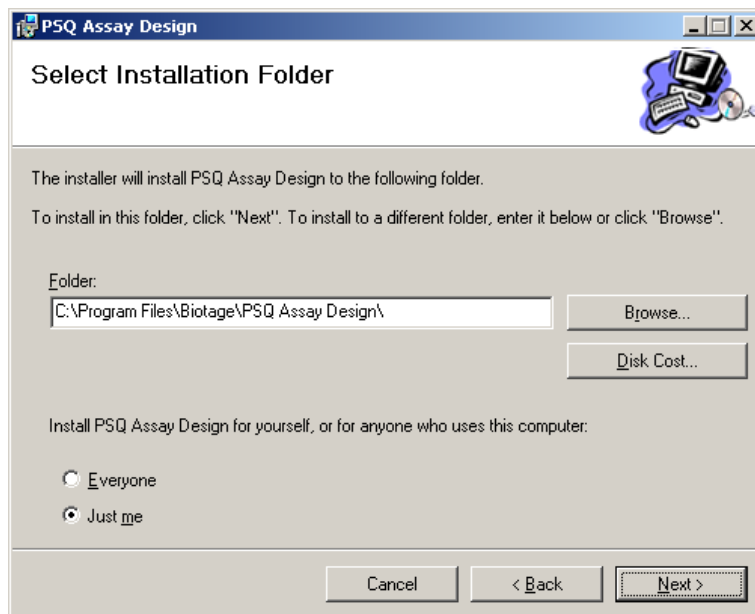


6. Click **I Agree** and **Next** to accept the terms in the license agreement.

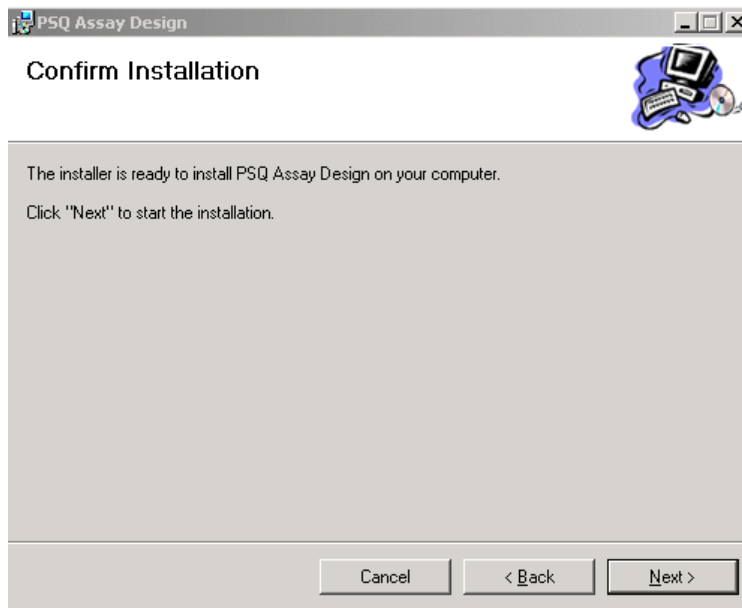


7. Choose the destination folder for the installation files. The default location is C:\Program Files\Biotage \PSQ Assay Design\. Click **Next** to proceed.

Note: If the **Just me** box is checked, the program will only be visible to the user who was logged in when the installation was performed.



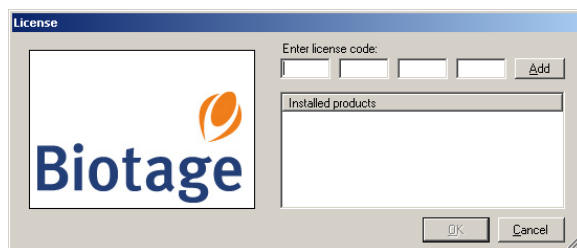
- Review the installation settings and click **Next** to start the installation.



- The installation wizard shows the progress of the installation.
- Click **Close** to exit the wizard.

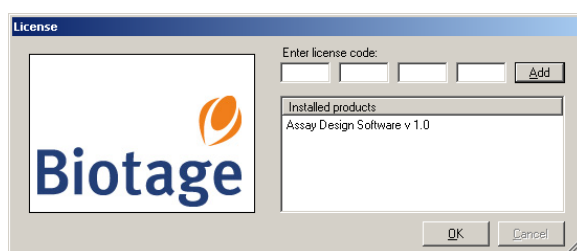
2.3 Starting Pyrosequencing Assay Design Software

- In the Windows **Start** menu, choose **Programs > Biotage > PSQ Assay Design**. Alternatively, double-click the **PSQ Assay Design** icon on the desktop.
- The first time the program is started, a **License** dialog opens.



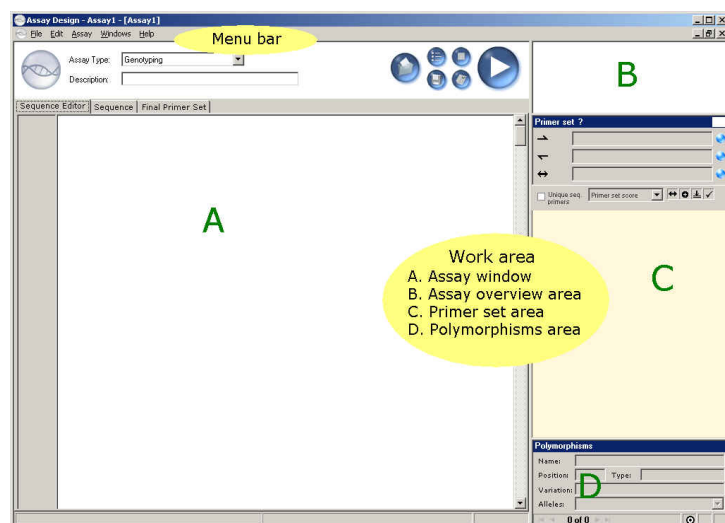
- The license key can be found on a slip of paper in the CD jacket. Enter the license key information in the **Enter license code** field and click **Add**. Assay Design Software v 1.0 will appear in the **Installed products** field.

Note: Every user (with an individual user account on the computer) will need to enter the license key the first time they start the program.



- Click **OK**. The Assay Design Software start screen will open.

2.4 The Assay Design Software start screen



Menu bar The menu bar contains 5 different drop-down menus: **File** -create, open and close assays, import sequences, or change default settings, **Edit** -cut and paste sequence information for analysis, or search for a sequence string **Assay** -analyze assays individually or in batch, view and change analysis settings, Zoom in/out on the sequence, **Windows** -hide/show the results panel and arrange windows, and **Help**.

Work area The work area displays the **Assay** window (A) and the results panel at start. The results panel is divided into three areas: the **Assay Overview** area (B), the **Primer set** area (C) and the **Polymorphisms** area (D). The **Assay** window and the results panel are used together to perform an assay design. For more information, see **Chapter 4 -Performing an assay design**.

Assay window This window can be used to choose assay type, enter or import the DNA sequence, set target and PCR primer regions (optional), run the assay design, view results, adjust the assay primers (optional), save the assay and view an assay report. For more information, see **Chapter 4**.

Assay overview area This area gives an overview of the entire sequence with symbols for polymorphisms, target regions, primers, and mispriming sites. This area can be used to get an overview of the whole sequence, including primer positions and mispriming sites. It can also be used for quick navigation to a specific part of the sequence on the **Sequence** tab in the **Assay** window. For more information, see **Chapter 4**.

Primer set area This area can be used to view a list of primer set candidates and to select the final primer set to be imported into PSQ system software. It is possible to edit and modify primers. Assay Design Software can also analyze previously designed primers by pasting or typing the sequences in to this area. For more information, see **Chapter 4**.

Polymorphisms area The polymorphisms or unknown sequences entered into the **Sequence editor** in the **Assay** window are automatically displayed in the **Polymorphisms** area. This area can be used to name the polymorphic positions (Position names), step between polymorphisms, and to define the target region for the next analysis. For more information, see **Chapter 4**.

3 Assay design settings

3.1 Change default assay design settings

1. Select **File | Default Settings...** in the menu bar. The **Default Assay Settings** dialog opens.
2. Change the settings as desired by entering new values, checking/unchecking the desired boxes and clicking **Set As Default**.
3. A message box with the following question appears.

"Do you want to apply your new settings to all open assays? If you choose No, the new settings will only be applied to new assays."

Click the button **Yes** to make the new default settings apply to currently open assays, as well as to all future, new assays. To view a description of the parameters that can be changed, see **section 3.3**.

3.2 Change assay design settings for one assay

1. In the Assay window, click the button **Settings for this assay**. The **Current Assay Settings** dialog opens where the settings for the active assay are set.
2. Change the settings as desired by entering new values, checking/unchecking the desired boxes and clicking **OK**. The settings are changed for this assay only (as long as the button **Set As Default** has not been clicked). To view a description of the parameters that can be changed, see **section 3.3**.

3.3 Description of settings and buttons

3.3.1 PCR Primer settings

Setting	Description	Factory settings	Optimal range
Min Primer Length (bp):	Minimum length of the PCR primers to be generated.	18 bp	18 - 24 bp
Max Primer Length (bp):	Maximum length of the PCR primers to be generated.	24 bp	
Optimal Amplicon Length from (bp):	Lower limit of the optimal amplicon range.	50 bp	50-250 bp
Optimal Amplicon Length to (bp):	Upper limit of the optimal amplicon range.	250 bp	
Max Amplicon Length (bp):	Maximum allowed length of the amplicon.	600 bp	
Allow Primer Over SNP:	Check the box to allow annealing over SNPs in the DNA sequence.	Unchecked	-
Melting Temperature Algorithm:	Choose algorithm for calculation of PCR primer T _m	Nearest Neighbor	-
Primer Concentration (μM):	Primer concentration used in the PCR reaction.	0.2 μM	-
Min Melting Temperature (°C):	Minimum melting temperature of PCR primers.	56.0 °C	68-74 °C
Max Melting Temperature (°C):	Maximum melting temperature of PCR primers.	86.0 °C	
Max Allowed T_m difference:	Maximum allowed T _m difference between forward and reverse PCR primers.	10.0 °C	0-2 °C
Max GC Difference (%)	Maximum allowed GC difference between the forward and reverse PCR primers and the amplicon.	30%	0-10 %

3.3.2 Sequencing primer settings

Setting	Description	Factory settings	Optimal range
Min Primer Length (bp):	Minimum sequencing primer length.	15 bp	15-20 bp
Max Primer Length (bp):	Maximum sequencing primer length.	20 bp	
Min Distance from target (bp):	Minimum distance between the sequencing primer and target region.	0 bp	0-3 bp
Max Distance from target (bp):	Maximum distance between the sequencing primer and target region.	3 bp	
Allow Primer Over SNP:	Check the box to allow annealing over SNPs. Note: Primers will not be allowed to anneal over an SNP within the 7 last nucleotides of its 3'-end.	Unchecked	-
Generate Forward Primers:	Check the box to generate forward sequencing primers.	Checked	-
Generate Reverse Primers:	Check the box to generate reverse sequencing primers.	Checked	-

3.3.3 Primer set settings

Setting	Description	Factory settings
Primer Set #:	Number of primer sets to be listed in the Primer set area.*	100

* The number of primer sets listed in the **Primer set** area is tightly linked to the number of sets that will be generated and analyzed. Thus, if the number of shown primers is lowered, fewer primer sets will also be generated, possibly decreasing the chance of finding a high scoring primer set.

3.3.4 Buttons

Button	Description
Get Factory	Click to load the factory settings.
Get Default settings	In the Current assay settings dialog: Click to load the latest default settings.
Set As Default	Click to save the entered settings as default settings.
OK	In the Current assay settings dialog: Click to apply the settings to the current assay.
Cancel	Click to cancel any changes made and close the dialog.

3.4 Selected settings will affect primer set scoring

Some settings will affect the primer set scores while others, while noted, will not impact the scoring. Furthermore, parameters that do affect scores will affect them differently. Many parameters have an allowed range, within which primers are generated but given a penalty (i.e. a decreased score), and an optimal range within which they are generated without penalty. A detailed description of how different settings will affect primer set scores is shown below.

3.4.1 PCR Primer score effects

Setting	Factory settings	Optimal range	Effect on primer set score
Min Primer Length (bp):	18 bp		Primers within the set min and max lengths are generated. This analysis step can result in warning messages for short primers, but will not affect primer set scores.
Max Primer Length (bp):	24 bp	18 - 24 bp	
Optimal Amplicon Length from (bp):	50 bp	50-250 bp	Amplicons within the set optimal range are generated without penalties. Amplicons outside the optimal range, but shorter than the max length, will be generated but given a penalty that increases linearly.
Optimal Amplicon Length to (bp):	250 bp		
Max Amplicon Length (bp):	600 bp		
Allow Primer Over SNP:	Unchecked	-	If the box is checked, a PCR primer is allowed to overlap up to two SNPs located within the 5'-end half of the PCR primer. No penalty is associated.

Primer Concentration (μM):	0.2 μM	-	Primer concentration will affect the T_m calculations.
Min Melting Temperature ($^{\circ}\text{C}$):	56.0 $^{\circ}\text{C}$	68-74 $^{\circ}\text{C}$	Primers with T_m in the middle of the set range, ± 3 $^{\circ}\text{C}$, will be generated without penalty (i.e. with factory settings, T_m between 68-74 $^{\circ}\text{C}$ will not get penalized). Primers outside the optimal T_m range, but within the set min and max temperatures, will be generated but given a penalty that increases linearly.
Max Melting Temperature ($^{\circ}\text{C}$):	86.0 $^{\circ}\text{C}$		
Max Allowed T_m difference:	10.0 $^{\circ}\text{C}$	0-2 $^{\circ}\text{C}$	T_m differences within the optimal range, 0-2 $^{\circ}\text{C}$, will be generated without penalty. T_m differences higher than 2 $^{\circ}\text{C}$, but lower than the set max value, will be generated but given a penalty that increases linearly with T_m difference.
Max GC Difference (%)	30%	0-10%	GC differences within the optimal range, 0-10%, will be generated without penalty. GC differences higher than 10%, but lower than the set max value, will be generated but given a penalty that increases linearly with GC difference.

3.4.2 Sequencing primer score effects

Setting	Factory settings	Optimal range	Effect on primer set score
Min Primer Length (bp):	15 bp	15-20 bp	Primers within the set min and max lengths are generated. This analysis step can result in warning messages for short primers, but will not affect primer set scores.
Max Primer Length (bp):	20 bp		
Min Distance from target (bp):	0 bp	0-3 bp	Primers within the set min and max distance from target are generated without penalties.
Max Distance from target (bp):	3 bp		
Allow Primer Over SNP:	Unchecked	-	If the box is checked, a sequencing primer is allowed to overlap up to two SNPs without penalty. Note: Primers will still not be allowed to anneal over an SNP within the 7 last nucleotides of its 3'-end.
Generate Forward Primers:	Checked	-	No penalty effect.
Generate Reverse Primers:	Checked	-	No penalty effect.

4 Performing an assay design

4.1 Introduction

Assay design can be performed for genotyping, allele quantification (AQ) and sequence analysis (SQA). A proposed workflow for performing an assay design is shown below.

Work flow:

- **Create an assay setup**

Choose assay type, enter or import a DNA sequence, define the target region and PCR primer regions (optional) and assign position names to the polymorphisms to be analyzed.

- **Run an automatic design**

Run the automatic assay design to create a number of different primer set candidates. It is also possible to re-use one or two primers from a previous assay and generate matching primers, or enter a complete primer set and have it analyzed by the software.

- **View results and select a final primer set**

The results of the assay design are displayed as primer sets with different scores and quality. It is possible to view detailed results for the different primer sets by double-clicking on a set. The primer set with the highest score is automatically selected as the final primer set (can be changed if desired). When importing an assay into PSQ system software, the information for the final primer set will automatically be transferred to the created Entry.

- **Optional: Adjust the assay**

It is possible to adjust the assay by changing the assay design settings.

- **Save the assay**

The assay can be saved as an xml format file (with the file extension xml).

Note: The xml format is not associated with the Assay Design Software application. Therefore, assay files must be opened from the Assay Design Software and not by double-clicking on the file in, for example, Windows Explorer.

Assay files can be imported into PSQ system software, automatically generating an SNP Entry for the assay in the PSQ database.

Note: Once a primer set is imported into the PSQ software, the created Entry cannot be edited. If editing is required, select menu choice **Duplicate Entry** in PSQ software to create an editable copy of the Entry.

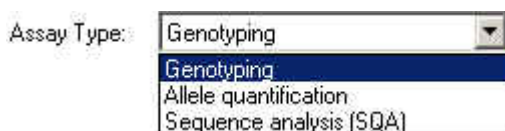
- **View an assay report of the selected primer set**

A report of the assay, containing detailed information on the selected primer set, can be viewed, saved and printed. Two different report formats are available: **Complete Results** and **Summary**.

4.2 Create an assay setup

4.2.1 Step 1 - Choose assay type and enter a description for the assay

1. After launching Assay Design Software, a new assay file is automatically displayed in the Assay window.
2. Choose assay type by selecting the desired assay (genotyping, allele quantification (AQ), or sequence analysis (SQA)) from the **Assay Type** drop-down menu. The analysis steps and primer scoring are automatically adjusted to the chosen application.



Note: A new **Assay** window will by default be the same **Assay Type** as the last assay that was analyzed.

3. If desired, enter a description for the assay in the **Description** field.

4.2.2 Step 2 - Enter the DNA sequence

1. Enter a DNA sequence on the **Sequence Editor** tab. There are four ways of entering a sequence (see **sections 4.2.2.1** and **4.2.2.2** for details):
 - Import a DNA sequence in GenBank-, EMBL- or FASTA-format.
 - Type in a DNA sequence
 - Copy and paste a DNA sequence
 - Open a previously saved assay file (*.xml)

Note: The entered sequence may not be longer than 10 000 characters.

The entered sequence is displayed on the **Sequence Editor** tab and on the **Sequence** tab of the **Assay** window.

2. If the sequence was typed or pasted into the **Sequence Editor**, position the cursor at the end of the sequence and click **Enter** to parse it, i.e. to get the sequence numbered and the nucleotides in the sequence divided up in blocks of ten. If the entered sequence contains characters that are not allowed, the first one will be highlighted in red. These characters must be corrected or removed before assay design can be performed.
3. On the **Sequence Editor** tab, the entered sequence is arranged in blocks of ten nucleotides. The polymorphisms are highlighted in bold and the position number for the first nucleotide of a row is shown in the column to the left of the sequence. The first polymorphism is automatically selected as target region (the target region can be changed if desired).

The sequence can be edited. As soon as the sequence is edited, the numbering and nucleotide grouping will disappear. Therefore, when editing has been finished, position the cursor at the end of the sequence and click **Enter** to parse the sequence again.

4. On the **Sequence** tab, the entered sequence is displayed together with its complementary strand. In addition to the sequence, a melting temperature (T_m) graph and Gibbs free energy (ΔG) graph of the template sequence are displayed. Default

target region and PCR regions can be viewed.

5. The **Assay overview** area gives an overview of the entire sequence with symbols for polymorphisms and target regions. It can be used to quickly navigate in the sequence on the **Sequence** tab.
6. Polymorphisms and/or target sequences for SQA (if unknown) are automatically displayed in the **Polymorphisms** area. It is possible to skip between polymorphisms and change target regions in the **Polymorphisms** area. It is also possible to select which allele to display in the sequence on the **Sequence Editor/Sequence** tabs. The different polymorphisms can be assigned **Position names**, which will then automatically be transferred to the Entry upon import into PSQ 96MA or PSQ HS 96A system software.

4.2.2.1 DNA sequence entry

- *Import a DNA sequence*

Select **File | Import Sequences...** or click on the button **Import sequence to this assay window** in the **Assay** window. The **Import Sequence** dialog is displayed.

Locate the sequence to import (*.txt) and click **Open**.

Sequences in GENBANK, EMBL and FASTA format can be imported. FASTA files can contain multiple sequences for simultaneous import. The different sequence formats are exemplified in **section 4.2.2.2**.

Note: To perform import of multiple sequences in FASTA format, **File | Import Sequences...** should be selected. The button **Import sequence to this assay window** only works for import of one sequence at a time.

By default, the software only shows files with the ending .txt. If the text file containing the sequence to import is not shown in the list because it has a file ending other than *.txt (e.g. if it originates from a Mac or UNIX system), select **All Files (*.*)** in the **File format** drop-down list to display all files in the folder.

- *Type in a DNA sequence*

Type the sequence as a continuous text in 5'-3' direction. The allowed characters for sequence input are **A, C, G,** and **T**, representing the four possible nucleotides adenosine, cytidine, guanosine, and thymidine, as well as IUPAC codes for polymorphisms (**W, R, K, Y, S, M, B, H, D, V, N**). Enter up to 10 000 nucleotides.

Polymorphisms (genotyping/AQ) and unknown regions (SQA) should be entered using the respective formats described in **sections 12.1 and 12.2**. Nested polymorphisms are not allowed.

- *Copy and paste a DNA sequence*

Copy a sequence from a text editor (e.g. Microsoft Word or Notepad) or from an Internet browser, and paste it into the empty area on the **Sequence Editor** tab.

If the sequence contains row numbers, these will automatically be removed when pasting the sequence into **Sequence Editor**.

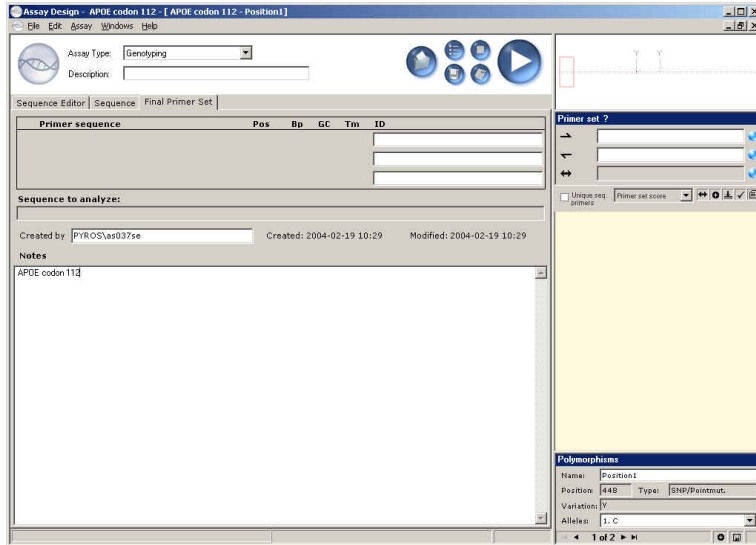
To copy a sequence from the **Sequence Editor** of one assay window, to the **Sequence Editor** of a second assay window, right-click on the sequence and select menu alternative **Copy Entered Sequence**. This is useful, for example, when designing assays for multiple SNPs within the same PCR amplicon.

- Open a previously saved assay file (*.xml)

Select **File|Open** to open a previously saved assay file (*.xml) in the **Assay** window.

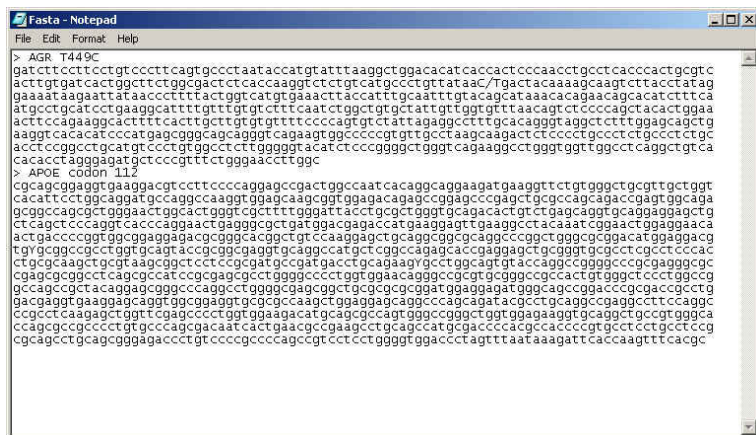
4.2.2.2 Import DNA sequences

FASTA file format: Using FASTA format, several sequences can be imported simultaneously into the Assay Design Software. For each sequence in the FASTA text file, one assay window will be opened automatically and the respective sequence imported into the **Sequence Editor** tab. The information in the header line preceding each sequence will be imported into the **Notes** field on the **Final Primer Set** tab, and will also be the default name of the assay file created.

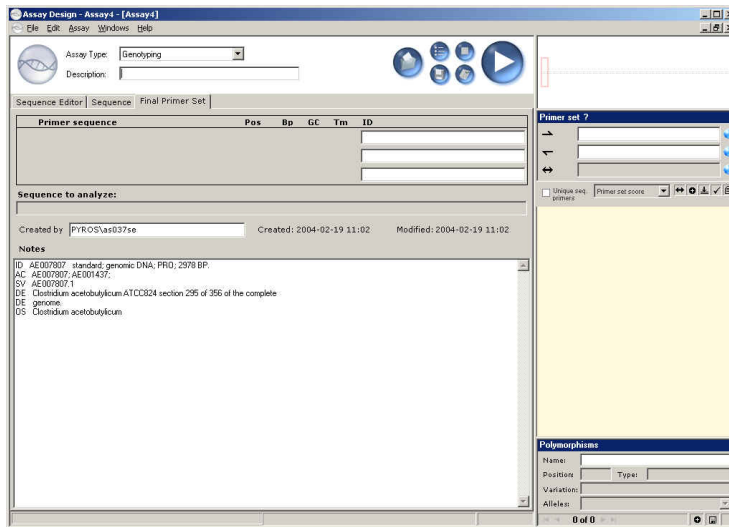


The FASTA file format is a common format for DNA sequence files. Sequences in FASTA file format are preceded by a line starting with the symbol ">" as the first character. The rest of the line is the name and description of the sequence (the header line). The following lines contain the sequence data. The sequences should not have any numbers (e.g. line numbers) and should contain a maximum of 10 000 characters each. SNPs should be denoted either with IUPAC codes, or with slash-notation (e.g. C/T). Insertion/Deletion polymorphisms should be typed in square brackets (e.g. [C]).

The sequences in FASTA format should be saved as text (*.txt) files. An example of a FASTA file containing multiple sequences is shown below.



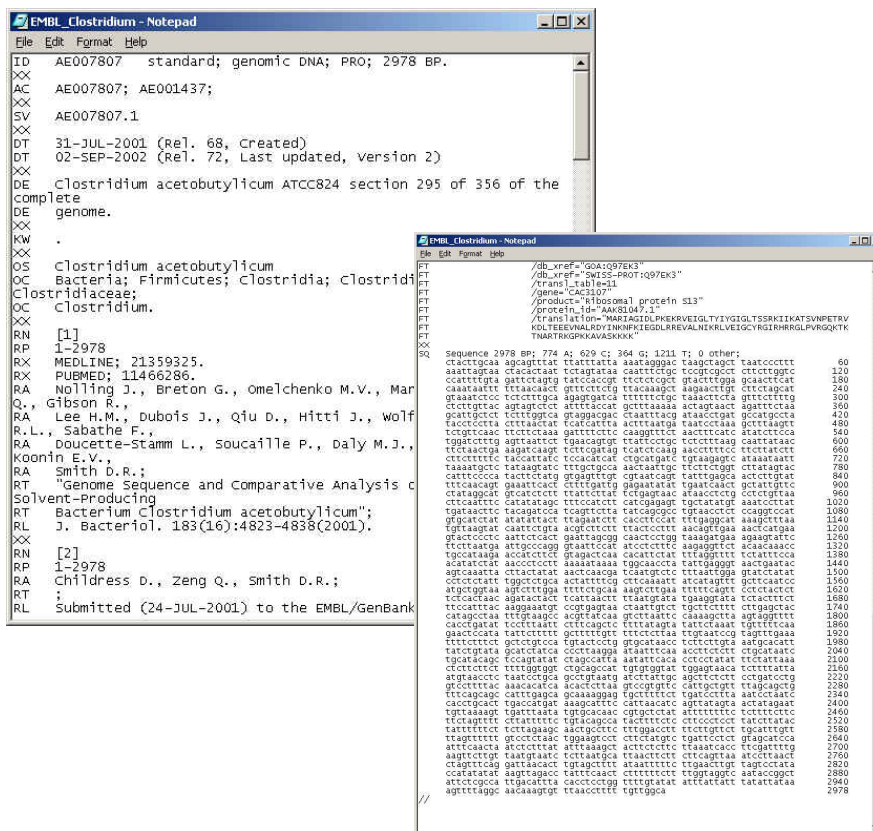
EMBL: Sequences in EMBL format, derived from the EMBL Nucleotide Sequence Database and saved in text format (*.txt), can be imported into the Assay Design Software one at a time. The information in the top rows of the EMBL file will automatically be transferred into the **Notes** field on the **Final Primer Set** tab (see example below).



Sequence files in EMBL format start with a number of information lines followed by the sequence. The sequence will contain line numbers. Only sequence files up to 10 000 characters in length can be imported. SNPs can be entered manually in the text file and should be denoted either with IUPAC codes, or with slash-notation (e.g. C/T). Insertion/Deletion polymorphisms should be typed in square brackets (e.g. [C]).

The sequence in EMBL format should be saved as a text (*.txt) file. The text file should start with the **ID** line. The end of the sequence, at the bottom of the text file, should be a double slash: **//**. For further information about how to save an EMBL sequence record on the correct text format, see **Save a file on EMBL text format**.

An example of an EMBL text file is shown below.

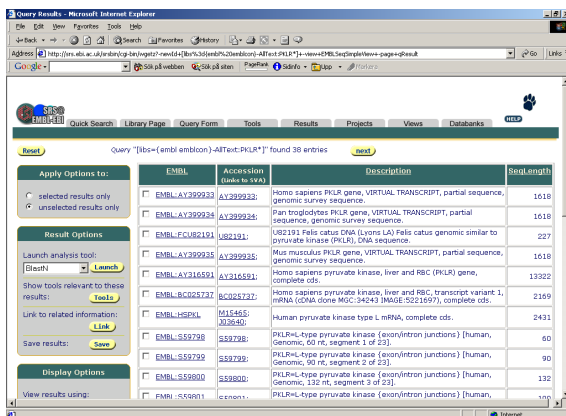


Save a file on EMBL text format

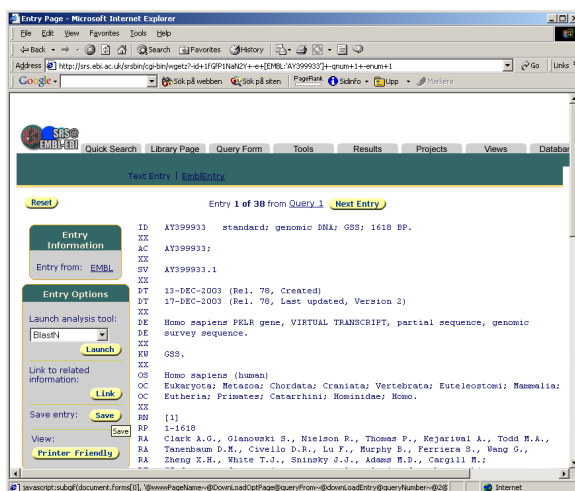
1. At the web site <http://www.ebi.ac.uk/>, the **Nucleotide sequences** database can be searched either for a gene name, or for a specific accession number. Type in the search item and hit the **Go** button.



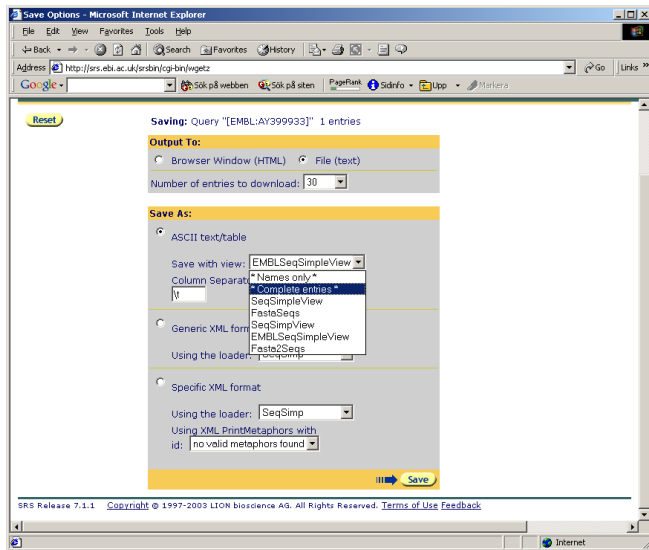
2. The search results in a number of hits. Click the EMBL number of the record of choice to open up a sequence record.



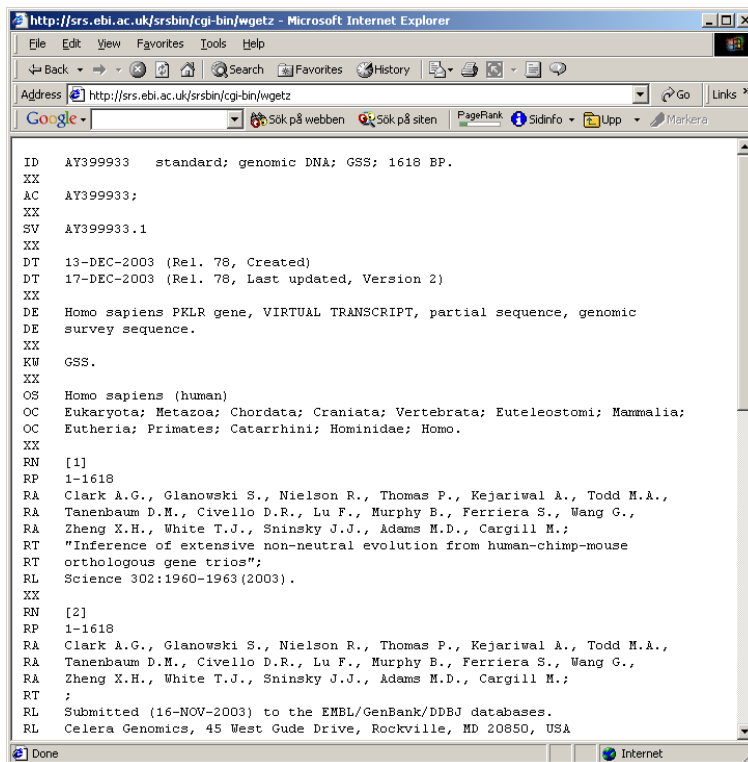
3. Click the **Save** button in the left panel of the sequence record to save the entry.



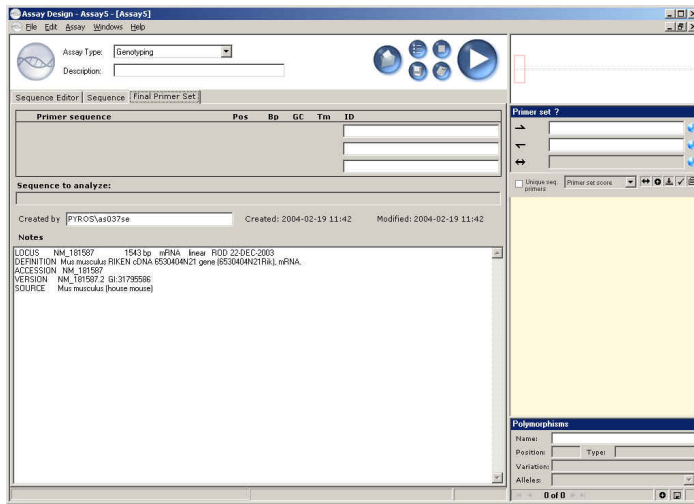
4. Change the view from EMBLSeqSimpleView to *Complete entries* and hit the **Save** button.



5. The Entry that opens is now in the correct format. Select **File | Save As...**, enter a file name of choice and change the file type to Text File (*.txt). Press the button **Save**.



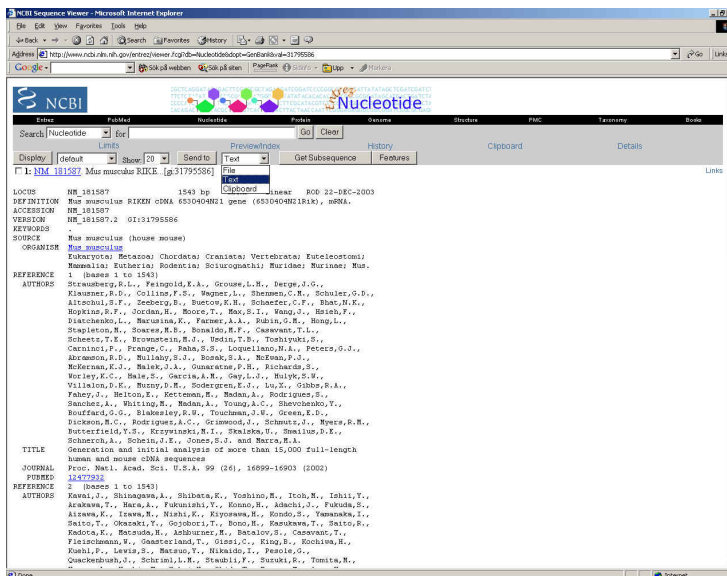
GenBank: Sequences in GenBank format, derived from the NCBI GenBank database and saved in text format (*.txt), can be imported in Assay Design Software one at a time. The information in the top rows of the GenBank file will automatically be transferred into the **Notes** field on the **Final Primer Set** tab (see example below).



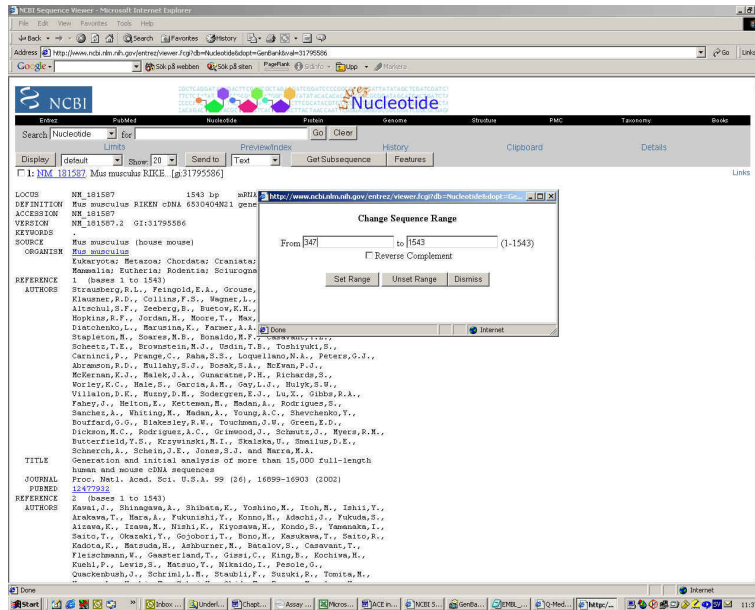
Sequence files in GenBank format start with a number of information lines followed by the sequence. The sequence will contain line numbers. Only sequences up to 10 000 characters in length can be imported. If the sequence in the database record is too long to import, shorten the sequence using GenBank functionality (See **Save a file on GenBank text format**).

Save a file on GenBank text format

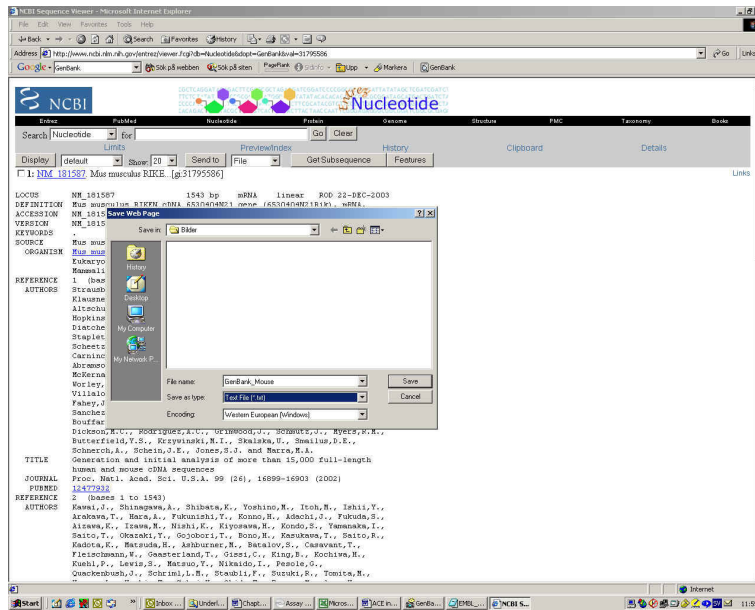
1. Change the format to **Text** and press the button **Send to**.



- If the sequence is too long to be imported into the Assay Design Software (i.e. more than 10000 characters), press the button **Get Subsequence** and select the desired sequence range in the dialog that appears.



- Select **File | Save As...**, enter a file name of choice and change the file type to Text File (*.txt). Press the button **Save**.



SNPs can be entered manually in the text file and should be denoted either with IUPAC codes, or with slash-notation (e.g. C/T). Insertion/Deletion polymorphisms should be typed in square brackets (e.g. [C]).

The sequence in GenBank format should be saved as a text (*.txt) file. The text file should start with the **LOCUS** line. The end of the sequence, at the bottom of the text file, should be a double slash: **//**. An example of a GenBank text file is shown below.

```

GenBank_3 - Notepad
File Edit Format Help
LOCUS NM_181587 1543 bp mRNA linear ROD 22-DEC-2003
DEFINITION Mus musculus RIKEN CDNA 6530404N21 gene (6530404N21R1k), mRNA.
ACCESSION NM_181587
VERSION NM_181587.2 GI:31795586
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 1543)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D.,
Aitschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M., Scheetz, T.E., Brownstein, M.J., Usdin,
Carninci, P., Prange, C., Raha, S.S., Lt.
Abramson, R.D., Mullay, S.J., Bosak, S.,
Mckernan, K.J., Malek, J.A., Gunaratne,
Worley, K.C., Hale, S., Garcia, A.M., Ga
Villalon, D.K., Muzny, D.M., Sodergren,
Fahey, J., Helton, E., Kettman, M., Mao
Sanchez, A., Whitting, M., Madan, A., You
Bouffard, G.G., Blakesley, R.W., Touchr
Dickson, M.C., Rodriguez, A.C., Grimwoc
Butterfield, Y.S., Krzywinski, M.I., Sk
Schmerchl, A., Schein, J.E., Jones, S.J.
Generation and initial analysis of mc
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26)
12477932
REFERENCE
AUTHORS 2 (bases 1 to 1543)
Kawai, J., Shinagawa, A., Shibata, K., Y
Arakawa, T., Haru, A., Fukunishi, Y., Kc
Aizawa, K., Izawa, M., Nishi, K., Kiyosa
Saito, T., Okazaki, Y., Gojobori, T., Bc
Kadota, K., Matsuda, H., Ashburner, M.,
Fleischmann, W., Gaasterland, T., Gissi
Kuehl, P., Lewis, S., Matsuo, Y., Nikita
Quackenbush, J., Schriml, L.M., Staubli
Wagner, L., Washio, T., Sakai, K., Okido
Baldarelli, R., Barsh, G., Blake, J., Bc
Carninci, P., de Bonaldo, M.F., Brownst
Fletcher, C., Fujita, M., Gariboldi, M.,
Hofmann, M., Hume, D.A., Kamiyama, M., Lee
Marchionni, L., Mashima, J., Mazarrelli
Ring, B., Ringwald, M., Rodriguez, I., S
Sato, K., Schonbach, C., Seya, T., Shiba
Toyo-oka, K., Wang, K.H., Weitz, C., whi
Wynshaw-Boris, A., Yoshida, K., Hasegaw
and Hayashizaki, Y.
TITLE Functional annotation of a full-lengt
JOURNAL Nature 409 (6821), 685-690 (2001)
PUBMED 11217851
REFERENCE
AUTHORS 3 (bases 1 to 1543)
Shibata, K., Itoh, M., Aizawa, K., Nagac

GenBank_3 - Notepad
File Edit Format Help
COMMENT PREDICTED REFSEQ: This mRNA record is supported by experimental
evidence; however, the coding sequence is predicted. The reference
sequence was derived from BC052355.1.
on Jun 17, 2003 this sequence version replaced gi:31712011.
FEATURES
source Location/Qualifiers
1..1543
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="19"
1..1543
/gene="6530404N21R1k"
/db_xref="GeneID:67795"
/db_xref="LocustID:67795"
/db_xref="MGI:1915045"
55..153
/gene="6530404N21R1k"
/product="RIKEN CDNA 6530404N21"
/protein_id="NP_853618.1"
/db_xref="GI:31712011"
/db_xref="GeneID:67795"
/db_xref="LocustID:67795"
/db_xref="MGI:1915045"
/translation="MSRVLVVGAGLGLSGLCAALLRKETAPLYLGL"
ORIGIN
1 0gggacggcg aggtctccgcg tctctgtcag cggcctagcc gaaggcagc cgcctatgcc
61 cgggtactgg tctgtggggcg tggcctaacc ggaagtctgt gtgccgcgct gctgaggaag
121 gaaataaccg ccccccctgta cctcggcctg tgagacaagg gtaggagatg aggggagaaga
181 atgattactg ccagcagfcc tctaatccc cagtgacacg ctgacttggg agctcagtac
241 atcacctgct ctctctatta tgtcaagag caccaaaatt tttatgaga actgttagct
301 catggaattt tgaagcctct gacatcccc attgaaggaa tgaaggaa gaaaggagat
361 tgcactttg tggcacctca agattttct tcagtatca agtactact gaaaaagtca
421 ggtcagaaag tcttccctca gactgtgtg actcagatc acctgaaaga taacaagtgg
481 gaagtctcca cagacaactg ctctgtgag cagttgacc ttgtatcct caccatgcc
541 gctcctcaga tctggaact tcaaggtgac attgtgaact taattagta acgccagagg
601 gagcaactga aatctgtgag ctactcctct cgtatgctc tggcctctt ttaagaata
661 gpcatgaaga ttgtgtgccc ttgtccctgc cgtacctca gcaagtcaacc ctgcatatgc
721 ttcacttcca ttgataataa gaagcgcaat atagagcat cagaatrgg tccatccgtg
781 gtgatccaaa ccactgtccc attgagatt caaacctgg aggccagtga ggcggatgtg
841 cagaagttaa tgcactcaga attggaacc attctgccgg gttgctca gccagttgct
901 acctatgac afaaatggac atattcacg gttacaagct cagtttcga cagacctgg
961 cagatgact ttcactcaa gcccttctct ggtgcggag gggatggatt tactcactcc
1021 aaactcaatg gctgcactc ctctccctg agtgcctga aggtttaaa gcgttatatt
1081 taagtctctg gttcttctc tctacattt ctttaaggatt tttgtttca caatcctctg
1141 ttattaattg tttctattt cttctctga agagaatta ctttgaaaa tgccttcact
1201 tactgttatt tttaatatg aatataaaa caatgttata attttgatg taccaccca
1261 tagtaaatg ataactctt aggtatatt tcttctcact ttttctgact ttattactaa
1321 tccttatgg agaaaaatg tgcttctaa tcacaataa cccaagccc gaattatact
1381 aaaccaaag tctcattcat ggaataact tatatcctt tgaaaacct cagttctaaa
1441 taataaacg ttaactaac aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa
1501 aaaaaaaaa aataaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa
//

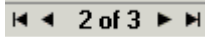
```

4.2.3 Step 3- Name Polymorphisms and unknown sequences

4.2.3.1 Name polymorphisms and unknown sequences in the Polymorphisms area

An unknown sequence is treated in the same way as a polymorphism in the **Polymorphisms** area.

1. Display the desired polymorphism to give a Position name by using the arrows at the bottom of the **Polymorphisms** area.



2. Enter a position name in the **Name** field. This is the position name that will be imported into the PSQ 96MA or PSQ HS 96A Entry.

4.2.4 Step 4- Optional: Set the target region

Genotyping/allele quantification: Assay Design Software automatically sets the first polymorphism in the entered sequence as the target region for Pyrosequencing analysis. This is indicated by a light blue highlight of the polymorphism on the **Sequence** tab and is shown in the title bar of the **Assay** window.

Sequence analysis: The last unknown region, including three known nucleotides flanking either side of the unknown sequence area, is automatically set as the target region for Pyrosequencing analysis by Assay Design Software. This is indicated by a light blue highlight of the nucleotides on the **Sequence** tab.

Note1: It is possible to select a target region that covers more than one polymorphism. If there are two polymorphisms in close proximity that are to be analyzed in the same Pyrosequencing reaction, the target region needs to cover both positions. Otherwise, the Sequence to analyze generated by the program will be cut to exclude the second polymorphism. Sequencing primers generated by the program will never overlap any part of the selected target region.

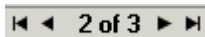
Note2: For selection of target region of repeat polymorphisms/STRs, see **Section 9.9** in **Hints & Tips** for special guidelines.


4.2.4.1 Change the target region

The target region is set either on the **Sequence** tab, on the **Sequence editor** tab, or from the **Polymorphisms** area. How to set the target region in the **Polymorphisms** area and on the **Sequence** Tab, respectively, is described below.

Skip between polymorphisms and set target region in the Polymorphisms area

1. If the entered sequence contains several polymorphisms, skip between them in **Sequence Editor** by clicking the arrows in the **Polymorphisms** area.



2. Once the desired target polymorphism or unknown sequence region has been selected, click the **Set Target** button .

Set the target region on the Sequence tab

1. **Genotyping/AQ:** On the **Sequence** tab, mark the polymorphism(s) to analyze and the desired sequence length before and after the polymorphism(s).

SQA: On the **Sequence** tab, mark the region to analyze. If analyzing unknown regions, include 2-3 known nucleotides flanking the unknown sequence in the target region.

Note: It is possible to search for a sequence motif to set as target region (see below).

2. Right-click on the marked sequence and select **Target Region | Set Target Region**. Alternatively, select **Assay | Target Region | Set Target Region** in the menu bar. The target region for Pyrosequencing analysis is highlighted in blue. On the **Sequence Editor** tab, the nucleotides in the selected target region are highlighted in yellow. Generated sequencing primers will be placed outside of this region.

3. The software automatically defines a PCR amplification region around the selected target region, which is used for PCR primer design.

Note: It is possible to manually redefine within which sequence region the forward and reverse PCR primers should be allowed to anneal. This is useful, for example, when the sequence contains multiple polymorphisms that should be contained within the same PCR amplicon. For further details, see **section 4.2.5**.

4. The assay setup is now complete. Continue with the instructions in **section 4.3**.

Search for a sequence motif in the entered sequence

1. Select **Edit | Find** (or use shortcut key **Ctrl+F**). The **Find** area is displayed at the bottom of the **Sequence/Sequence editor** tab.

2. Enter the sequence motif to search for in the **Find what** field.

Note: It is not possible to search for a polymorphism (e.g. C/T) or a sequence string containing a polymorphism. The search string should only contain the characters A, C, G, or T.

3. There are four different search directions to choose from.

- Forward sequence – search for a motif from left to right on the upper strand.
- Reverse sequence – search for a motif from left to right on the lower strand.
- Forward complementary sequence - search for a motif from right to left on the upper strand.
- Reverse complementary sequence - search for a motif from right to left on the lower strand.

4. Click **Next** to find the next occurrence of the motif. If found, it is marked on the **Sequence** and **Sequence Editor** tabs. Click **Previous** to search for the former occurrence of the motif.

5. Click the **cross** button  to close the **Find** area.

4.2.5 Step 5 - Optional: Redefine PCR primer regions

The software automatically defines PCR priming regions around the target, which are used for the PCR primer design. It is possible to manually redefine within which sequence region the forward and reverse PCR primers should be allowed to anneal.

Note: If there are two or more polymorphisms that should be contained within the same PCR amplicon, the PCR primer regions need to be manually defined. Select the PCR primer regions so that the forward PCR primer is only allowed to anneal upstream of the first SNP, and the reverse PCR primer is only allowed to anneal downstream of the last SNP.

Define forward and reverse PCR primer regions

1. On the **Sequence** tab or **Sequence Editor** tab, mark the part of the sequence within which the forward PCR primer should be generated.
2. Right-click on the marked sequence and select **PCR Regions | Set Forward PCR Primer Region**.
3. Mark the part of the sequence within which the reverse PCR primer should be generated.
4. Right-click on the marked sequence and select **PCR Regions | Set Reverse PCR Primer Region**.
5. Bars placed over the sequence on the **Sequence** tab indicate the selected regions. On the **Sequence Editor** tab, the selected regions turn blue. To hide the bars on the **Sequence** tab, right-click on the sequence and deselect **View | PCR regions**.
6. The assay setup is now complete. Continue with the instructions in **section 4.3**.


4.3 Run an automatic design

4.3.1 Generate new primer sets

1. Click the **Run Assay Design** button to perform the assay design. The progress of the analysis is shown in the form of a progress bar at the bottom of the work area.



While analyzing, the button displays two rotating arrows.

2.  It is possible to stop an ongoing assay design by clicking the red **Stop** button, or by using shortcut keys **Ctrl + Q**.
3. The resulting primer sets are displayed in the **Primer Set** area.



It is also possible to enter primer(s) and generate matching primers, resulting in primer sets with scores for Pyrosequencing analysis. Furthermore, a complete primer set can be entered and analyzed to obtain a score and quality ranking. See **section 4.3.2**.

4.3.2 Analyze previously designed primers



Re-use one or two primers from a previous assay and generate matching primers



1.
 - a. On the **Sequence** tab, mark where the primer anneals to the DNA sequence. The **Find** function can be used to search for a primer sequence string in forward or complement/reverse direction (see **section 4.2.4.1**).
 - b. Right-click and select one of the following:
 - **Set As PCR Primer | Biotinylated** if it is a PCR primer that should be biotinylated (it will be set as forward or reverse depending on which side of the target region the marked sequence is situated)
 - **Set As PCR Primer | Not biotinylated** if it is a PCR primer that should not be biotinylated (it will be set as forward or reverse depending on which side of the target region the marked sequence is situated)
 - **Set As Sequencing primer** if it is a sequencing primer

The primer appears on the **Sequence** tab and at the top of the **Primer set** area in the correct field (forward/reverse PCR primer field or sequencing primer field).

Alternative:

- a. At the top of the **Primer set** area, type or copy-and-paste the primer sequence in the appropriate field (forward/reverse PCR primer field or sequencing primer field).

Note: Before a sequencing primer can be entered in the sequencing primer field, its direction must be defined. This is done by pressing the **Set sequencing primer direction** button  and selecting which PCR primer should be biotinylated.
 - b. As soon as you leave the primer entry field, the software automatically places the primer at its annealing site in the DNA sequence on the **Sequence** tab. Check that this is the correct annealing site.
2. If desired, repeat the above procedure for the second primer (for example, two PCR primers can be set and matching sequencing primers found by the software).
 3. The entered primers are automatically locked as soon as you leave the primer entry field. This is indicated by a darkened blue button  to the right of the respective primer in the **Primer set** area.



By locking the primer(s), the software keeps this primer(s) constant and tries to find the best primer(s) complementing the locked primer(s). To toggle between locked/unlocked primers, click the blue button to the right of the primers. Darkened buttons  indicate locked primers and highlighted buttons  unlocked primers.

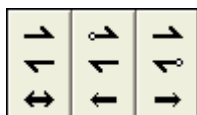
- Click the **Run Assay Design** button to start the assay design.



- The results (primer sets containing the locked primers in combination with matching primers) are listed in the list in the **Primer set** area.

Enter a complete primer set and obtain a score for Pyrosequencing analysis

- If a new primer set is to be added to a list of generated primers, click the **New primer set** button . The fields at the top of the **Primer set** area are emptied to allow entry of new primer sequences.
- Click the **Set sequencing primer direction** button  in the **Primer set** area. The following primer set options appear:




- Biotinylation not defined
- Biotinylation of forward PCR primer (reverse assay)
- Biotinylation of reverse PCR primer (forward assay)

- Select the desired primer set combination. The fields in the **Primer set** area are cleared and new primers can be entered.
- At the top of the **Primer set** area, type or copy-and-paste the three primers in the appropriate fields (forward/reverse PCR primer field and sequencing primer field).
 - The software automatically places the primers at their respective annealing sites on the DNA sequence on the **Sequence** tab. Check that these are the correct annealing sites.



Alternative:


- On the **Sequence** tab, mark where the primer anneals to the template sequence. The **Find** function can be used to search for a primer sequence string in forward or complement/reverse direction (see **section 4.2.4.1**).
- For each of the three primers, right-click and select one of the following:
 - Set As PCR Primer | Biotinylated** if it is a PCR primer that should be biotinylated (it will be set as forward or reverse depending on which side of the target region the marked sequence is situated)
 - Set As PCR Primer | Not biotinylated** if it is a PCR primer that should not be biotinylated (it will be set as forward or reverse depending on which side of the target region the marked sequence is situated)
 - Set As Sequencing primer** if it is a sequencing primer

The primers appear on the **Sequence** tab and in their respective fields in the **Primer set** area (forward/reverse PCR primer field and sequencing primer field).

- All three primers are automatically locked as indicated by darkened buttons  to the right of each primer in the **Primer set** area. A score for the primer set will automatically be

generated as soon as you leave the primer entry fields.

Note: By locking the primer(s), the software keeps this primer(s) constant and tries to find the best primer(s) complementing the locked primer(s). To toggle between locked/unlocked primers, click the blue button to the right of the primers. Darkened buttons  indicate locked primers and highlighted buttons  unlocked primers.

6. Click the **Save candidate** button  to save the primer set to the primer set list. An asterisk below the primer ID indicates that it is a manually entered primer set.

4.4 View results

4.4.1 View results - overview

The results of assay design are displayed in the **Primer set** area in the form of a primer set list. Every primer set has been assigned a score and quality, which reflects its suitability for both PCR amplification and Pyrosequencing analysis. By default, the primer sets are sorted by primer set score (0-100, where 100 is the best score) so that the best primer set ends up at the top of the list.

The primer set list can be sorted in other ways, see **section 4.4.3**. One hundred primer sets are shown by default. This value can be changed in the **Default assay settings** dialog, if desired.

By default, the top score primer set is:

- shown at the top of the **Primer set** area
- displayed first in the primer set list and highlighted in light blue to indicate that it is selected
- defined as the final primer set, indicated by a dark gray box surrounding it (the selection of the final primer set can be changed if desired, see **section 4.4.4**).

4.4.2 Scoring and quality

Primer set description

#3	▷ F1	ACAGCATAGACAGATAGGACA
	◁ R3	TCTCTTGGTGTGCTCTAATTC
58	→ S1	GGACAGATATAGACACCAGA

The color in the left panel indicates the quality of the primer set. The arrows indicate the direction of the corresponding primer. The ring on the reverse PCR primer arrow indicates that this primer is biotinylated.

Color	Score	Quality and description
Blue	≥ 88	High quality.
Yellow	60-87	Medium quality.
Orange	40-59	Low quality.
Red	0-39	Bad quality. Discard the primer set.

Symbol	Description
#3	Primer set ID.
58	Primer set score.
F1	Forward PCR primer ID.
R3	Reverse PCR primer ID.
S1	Sequencing primer ID.

For a primer set with **high** quality, i.e. labeled with color code blue and with a score higher than 87, none of the analysis steps performed have identified any problems of concern. The software algorithms are very stringent in their analyses and a high score primer set can therefore be used directly without any further manual checks or analyses.

For a primer set with quality **medium** or **low**, i.e. labeled with color-codes yellow or orange, one or several of the analysis steps have identified problems that may be of concern. The software algorithms are quite stringent, and most medium score assays can be expected to work very well, whereas low score assays only should be used after very careful consideration. In general, two primer sets with similar scores are of equal quality and either one of them can be selected for Pyrosequencing analysis. For primer sets with equal scores it may sometimes be informative to see if the primer set score has been lowered because of a rather severe penalty in one analysis step or because of several, individually lower penalties. In this case, the latter primer set with many small penalties, rather than the first one with a single severe structure, will probably be the better choice.

A potentially severe structure in an analysis step is indicated by a Penalty > 50, which will generate a warning, visible in the information field on the **Sequence** tab. The higher the penalty, the larger is the risk of problems. If an individual analysis step gets Penalty 100, the problem is considered serious enough to set the whole primer set score to 0 and thereby make it **Discard** quality. Penalties lower than 50 are in general nothing to be concerned about. To view detailed information on different primer sets, see **section 4.4.3** for further instructions.

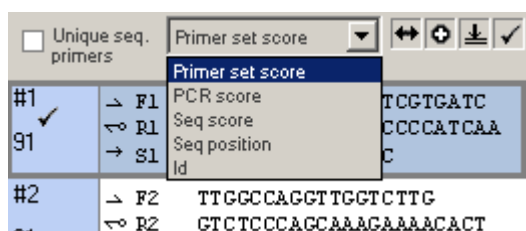
4.4.3 Sort the primer set list

In addition to sorting the results by primer set score, it is possible to sort the results by *PCR Score*, *Seq Score* and *Seq Position* and *ID*. This can be useful if the sorting of the primer set list by primer set score gives a low variability among the primer sets. For example, if you do not find enough different forward PCR primers to be displayed in the list.

The option to display **Unique** sequencing primers may be useful when performing multiplex assay designs.

To sort the primer set list:

1. Select the desired sorting option from the drop-down list in the **Primer set** area.



2. Check the **Unique seq. primers** box to display unique sequencing primers. Only the best primer set generated for each sequencing primer will be displayed in the list. This mode of display may be useful when designing multiplex assays.

4.4.4 View information on different primer sets

By default, the primer set with the highest score is selected as the final primer set by the software. A dark gray box surrounding the primer set indicates that it is selected as final. For a primer set with high quality, i.e. labeled with color code blue, none of the performed analysis steps have identified any problems of concern. Thus, a high score primer set can be used directly without any more checks or analyses.

For lower quality primer sets, however, a manual analysis may be required before use. To view more information on primer sets, follow the instructions below.

To view information on a primer set:

1. Click on the desired primer set in the **Primer set** area. The selected primer set is highlighted in light blue and displayed in the top fields of the **Primer set** area.


In the right-click menu, there is an option to **Copy Primer Set**. A copied primer set can be pasted into the entry fields of a different assay window, or into a text editor like Microsoft Word. There is also a menu alternative for copying the whole list, **Copy All Primer Sets**, to e.g. a Microsoft Word document.
2. On the **Sequence** tab, the primers are displayed together with the DNA sequence, as well as the graphs of the melting temperature and Gibb's free energy. The biotinylated PCR primer is marked with a ring at the 5'-end. Any warnings generated for the primers are indicated by a warning triangle next to the primer. When a new primer set is selected in the **Primer set** area, the graph on the **Sequence** tab is automatically updated.
3. The **Assay overview** area, at the top right corner of the work area, gives an overview of the entire sequence with symbols for polymorphisms, target regions, primers, and mispriming sites. Use this area to get an overview of the whole sequence, of primer positions and mispriming sites. Also use it to quickly navigate to a specific part of the sequence on the **Sequence** tab in the **Assay** window.
4. In the information area (on the **Sequence** tab), general information about the primers in the selected primer set, including primer sequence, length, warnings, sequence to analyze etc., can be viewed. If at least one analysis step has identified a problem of potential concern, a warning triangle and an associated warning message will be displayed in the information area. Use this as a quick indication of what analysis steps you need to check in the detailed report.

When a new primer set is selected in the **Primer set** area, the information area is automatically updated.

5. To view even more detailed information on the highlighted primer set, double-click the primer set **or** choose menu alternative **Assay | View Report...** **or** click the button **View assay report** in the **Assay** window **or** right-click on the primer set and select **View Report**. The **Report** window opens displaying detailed information on the primer set, e.g. the different analyses performed for each primer and primer combinations.

4.4.5 Optional: Select a different primer set as final

The final primer set is the set that will be imported into PSQ 96MA or PSQ HS 96A system software.

1. Right-click on the primer set of choice and choose **Select as final**. Alternatively, select the primer set and click the **Select as final** button . A surrounding dark gray box marks the chosen primer set.
2. The final primer set is displayed on the **Final Primer Set** tab of the **Assay** window.

On the **Final Primer Set** tab, it is possible to change the IDs of the primers, edit the creator of the assay and add notes about the assay.



4.5 Adjust an assay

4.5.1 Change the assay design settings

If the software fails to generate acceptable primer sets using the default settings, the results may be improved by changing the settings for PCR primer and sequencing primer design.

It is also possible to change the default settings to be applied on all new assays. See **Chapter 3**.

To change assay settings and re-analyze the assay:

1.  Click the **Settings for this assay** button in the **Assay** window. The **Current assay settings** dialog opens.
2. Change the settings by entering the desired values, checking/unchecking the desired boxes and clicking **OK**. The selected settings will only be used for the assay in which the **Settings for this assay** button was clicked. To view a description of the parameters that can be changed, see **section 3.3**.
3.  Click the **Run Assay Design** button to re-analyze the assay and generate primer sets based on the new settings.
4. The results are displayed in the primer set list in the **Primer set** area. View the results and select a final primer set, see **sections 4.4.3 and 4.4.4**, or try to adjust the assay settings in a different way.

Buttons in the Current assay settings dialog



Button	Description
Get Factory	Click to get the factory settings.
Get Default	Click to get the default settings.
Set as Default	Click to set the current settings as default settings and use the settings in the assay. The settings will also be applied in the Default Assay Settings dialog.
OK	Click to apply the selected settings for the assay in which the Settings button was clicked.
Cancel	Click to cancel any changes made and close the dialog.

4.5.2 Edit the PCR and sequencing primers

It is possible to edit PCR and sequencing primers, both on the **Sequence** tab and in the **Primer set** area. The edited primers will automatically be reanalyzed to generate a new primer set score. Editing primers can be useful e.g. when generating an assay with a universal PCR primer tail, or when adding nucleotides to the 5'-end of a PCR primer to circumvent a template loop.

For more hints & tips when editing primers, see **Chapter 9 - Hints & Tips**.

To edit PCR and sequencing primers:

1. If the primers to edit are locked (indicated by darkened blue buttons ) , click on the buttons to unlock them (unlocked buttons are highlighted ).

2. **a. Drag the primer along the sequence or change its length**


On the **Sequence** tab, click on the primer to edit and drag the primer along the template sequence to change its position. If the mouse-pointer is instead pointed at the end of the primer, the pointer icon will change to a double-arrow symbol. Dragging will then change the primer length, rather than moving the primer around.

The sequence of the primer, displayed in the **Primer set** area, is automatically adjusted to the template sequence to which it anneals. A new, re-calculated score and quality color are displayed at the top right corner of the **Primer set** area.

- b. Add/remove nucleotides in the primer sequence**

In the **Primer set** area, click in the desired primer field and enter/delete the desired nucleotides in the primer sequence. Only A, C, G, or T can be added to the primer sequence. A new, re-calculated score and quality color are displayed at the top right corner of the **Primer set** area as soon as you leave the primer field.

Note: PCR primers with universal tails, or primers with non-specific tails to avoid template loop formation, can be created in this way. See **Chapter 9 - Hints & Tips** for more information when editing primers.

3. A new primer set has been created. The primers are automatically re-analyzed and the score updated.
4. When satisfied with the results, click the **Save candidate** icon  in the **Primer set** area to save the primer set to the list. The primer set will receive a unique primer set ID and be placed at the bottom of the list. A star will indicate that it has been manually

added to the list.

View the result and select a final primer set, see **sections 4.4.3 and 4.4.4**, or try to adjust the assay in a different way.

4.5.3 Select one or more primers from a primer set and re-analyze the assay

It is possible to lock one or two of the primers in a primer set and re-run the analysis to generate different candidates complementing the one or two locked primers.

Example 1:

Select a certain sequencing primer and then generate a list of primer sets with PCR primers matching the sequencing primer.

Example 2:

Select the forward PCR primers and generate a list of primer sets with a matching reverse PCR primer and sequencing primer.

To lock a primer and generate matching pairs:

1. Click on the blue buttons to the right of the desired primers to lock them (the blue buttons will darken).
2. Click the **Run Assay Design** button in the Assay window to re-analyze the assay.



3. The results are displayed in the **Primer set** area. View the results and select a final primer set (optional), see **sections 4.4.2 and 4.4.3**, or try to adjust the assay in a different way.

4.6 Save an assay

An assay can be saved in .xml file format. The stored assay file can be used either to re-open the assay in Assay Design Software, or to import an Entry in PSQ 96MA or PSQ HS 96A system software.

To save the assay:

1. Select **File | Save** (or **Ctrl+S**) in the menu bar or click the **Save** button in the **Assay** window. The **Save as** dialog opens.
2. Locate the folder in which to save the assay.
3. Enter a name for the assay and click **Save**.



If changes are made to the assay and the assay is saved, the assay file is overwritten with the new information. To save the assay in a different folder and/or with another name, use the **Save As** command.

Note: The xml format is not associated with the Assay Design Software application. Therefore, assay files must be opened from the Assay Design Software and not by double-clicking on the file in, for example, Windows Explorer.

To save the assay in a different folder and/or with another name:

1. In the menu bar, select **File | Save As**. The **Save as** dialog opens.
2. Locate the folder in which to save the assay.
3. Enter a name for the assay and click **Save**.

4.7 View an assay report

There are several ways to display a report with detailed information for a selected primer set. The report can be printed and/or saved in either Html- or text-format.

4.7.1 View a report

To open the **Report** window, select the primer set by clicking on it in the **Primer set** area and do one of the following:

- double-click the primer set
- choose menu alternative **Assay | View Report...**
- click the button **View assay report** in the **Assay** window
- right-click on the primer set and select **View Report**
- press the short cut key combination **Ctrl + R**

The **Report** window opens displaying detailed information on the primer set, e.g. the different analyses performed for each primer and primer combinations. There are four different report formats to choose from as listed in the left part of the **Report** window.

Report formats

Report format	Description
Complete Results	A report in Html format. The report contains detailed information about the assay, and the selected primer set in particular. It presents details of the analysis steps that have been performed and any penalties and/or warnings that have been generated during the design.
Complete Results (Text)	A report in text format. The report contains detailed information about the assay, and the selected primer set in particular. It presents details of the analysis steps that have been performed and any penalties and/or warnings that have been generated during the design.
Summary	A short summary in Html format. The summary contains information about primer sequences and biotinylation, in a format that is suitable for ordering oligonucleotides.
Summary (Text)	A short summary in text format. The summary contains information about primer sequences and biotinylation, in a format that is suitable for ordering oligonucleotides.

4.7.2 Print a report

1. Choose which report format to use by clicking on the desired report format in the column to the left. If Complete Results (Text) or Summary (Text) are selected, a report in text format can be printed. Otherwise the report will be in HTML format.
2. Click the **Print report** button in the **Report** window.



Click the **Print preview** button in the **Report** window to preview the printout. Click **Close** to close the preview.



3. The standard **Print** dialog opens. Select the printer on which to print the report and click **Print**.

4.7.3 Save a report

A report can be saved in text- or html-format.

1. Choose which report format to use by clicking on the desired report format in the column to the left. If Complete Results (Text) or Summary (Text) are selected, a report in text format can be saved. Otherwise the report will be saved in HTML format.
2. In the **Report** window, click the **Save report** button.



The **Save as** dialog opens.

3. Locate the folder in which to save the report.
4. Enter a name for the report and click **Save**.

5 Performing batch assay design

5.1 Introduction

Automatic batch assay design on several sequences can be performed for genotyping, allele quantification and sequence analysis. When running batch analysis, all open assay windows will be analyzed with their respective settings.

A proposed workflow for performing a batch assay design is shown below.

- **Create assay setups**

Enter or import DNA sequences into different assay windows, choose assay type, define the target region and PCR primer regions (optional), and assign position names to the polymorphisms to be analyzed. See **Chapter 4 -Performing an assay design** for further details.

- **Run automatic batch assay design**

Run batch assay design for simultaneous analysis and design of all open assays. Start by defining the folder where the resulting assay files should be stored. During batch analysis, the program will be locked and cannot be accessed again until all open assays have been analyzed (or until batch analysis has been stopped).

- **Save the results**

After batch assay design has finished, the successfully analyzed assays will be automatically saved as xml format files (with the file extension xml) in a folder of choice, together with a text report of the best primer set per assay. Only failed assay files will remain open in Assay Design Software for evaluation and further analysis.

- **Optional: Adjust the assays**

If an assay is to be adjusted, open the relevant assay file and edit primers or change assay settings. See **Chapter 4 -Performing an assay design** for further details.

- **View an assay report of the selected primer set**

To view a detailed report of a designed assay, open the relevant assay file and view the report. See **Chapter 4 -Performing an assay design** for further details.

5.2 Run automatic batch assay design

Once target regions, position names and assay types have been selected for all open assays, batch analysis can be started.

1. Select **Assay | Setup Batch Assay Design**. This opens a dialog where the output directory for the resulting assay files and text report can be selected.
2. Select **Assay | Run Batch Assay Design** to open the **Batch Design Progress** dialog. Batch design is automatically initiated for all open assay windows.
3. Optional: The analysis can be stopped at any time by clicking the **Stop** button.
4. After analysis, all successful assays have automatically been saved and closed. Only failed assays will remain open in Assay Design Software. The output field of the Batch Design Progress dialog shows a short report with the best primer set generated per assay. This report can also be saved to the output folder, under the name PSQ Assay Design Log.txt, by clicking the **Save** button.

Note: If a new batch design is started, using the same output folder as for the first one, the new primer set information will be added to the bottom of the PSQ Assay Design Log.txt file. Thus, PSQ Assay Design Log.txt will never be over-written by consecutive analyses.

5. Close the Batch Design Progress dialog to regain access to Assay Design Software.

6 Importing an assay file into PSQ system software

The instructions below describe how to import assay files from within PSQ 96MA Software or PSQ HS 96A Software. Import will automatically create an SNP Entry for the imported assay file.

Note 1: Import of assays is only supported by PSQ 96MA version 2.1 or higher, and PSQ HS 96A version 1.2 or higher.

Note 2: Assays can only be imported into **SNP Simplex Entries**. **SQA Entries** and **SNP Multiplex Entries** do not support import of assay files.

To import an assay file into PSQ 96MA or PSQ HS 96A Software:

1. Open PSQ 96MA Software or PSQ HS 96A Software.
2. Select **SNP | Simplex Entries** to display the **Simplex Entries** tree view.
3. Right-click on the folder in which to save the assay file (the Entry) to import and select **Import Entries** from the pop-up menu. The **Entry Import** dialog opens.
4. Click **Browse**, locate and select the assay file (*.xml) to import and click **Open**. Alternatively, enter the file path to the file in the **File to import** field and click **Enter** on the keyboard. The Entry in the selected file is shown in the **Entries** area of the dialog.
5. At import, a dispensation order will be generated for the assay. By default, if this dispensation order would generate warnings in the PSQ system software, the import will be stopped and failed. If you want to override this, check the **Import if dispensation order gives warnings** box to import entries even if warnings are generated by the dispensation order algorithm.
6. Click **Import** to import the assay file (the Entry). During import, the software calculates a dispensation order, and warnings and error messages may also be generated. If there are many polymorphic positions in the sequence to analyze, the dispensation order generation may take time.
7. When import is complete, a dialog appears saying that the import has finished. It is possible to stop the import by clicking **Stop**.

Note: It may take some time before the import is stopped because the dispensation order generation cannot be interrupted.

8. The Entry in the **Entries** area is updated and shows the import status of the Entry (**Status** column) and, if generated, warnings (**Dispensation order warnings** column) and error messages (**Dispensation order errors** column). The status shows whether the Entry was imported or not and the reason if it was not imported. An icon in front of the Entry also indicates the status. Double-click on an entry in the **Entry ID** column to display an overview of the information in a window that opens above the list.

Successfully imported entries can now be opened from the simplex entries tree view and used in run setups.

Note: If import fails because an Entry with the same name already exists in the PSQ database, refer to **Section 9.10 Troubleshooting guide** for help.

7 Guidelines for PCR and sequencing primer design

7.1 Introduction

Assay Design Software generates primers that fulfill the specific requirements of Pyrosequencing analysis. Depending on the chosen assay type, the software carries out a number of analysis steps. Any potential problems detected in these analysis steps will generate penalties. The weighted sum of the penalties is used to calculate a total score for each primer set, and the result is a list of primer set candidates with different scores for quality and suitability for PCR and Pyrosequencing analysis.

A number of parameters are taken into account when using Assay Design Software to generate primer sets (PCR and sequencing primers). Many parameters, e.g. target distance, can be changed in the software using the **Default assay settings** dialog. The software always tries to achieve the set optimal conditions when designing the primer sets. Primer sets where the conditions deviate significantly from the set optimal conditions will get a higher penalty, and thereby a lower total score and quality.

The primer design part of the User Manual includes information on:

- The analysis steps carried out by the software, i.e. which parameters and secondary structures are checked and detected by the software when designing PCR primers and sequencing primers.
- Guidelines for PCR primer design.
- Guidelines for sequencing primer design.

7.2 Analysis steps performed by the software

7.2.1 Introduction

When designing PCR and sequencing primers for Pyrosequencing analysis in Assay Design Software, a number of analyses are carried out resulting in a list of primer set candidates with different scores (100-0, where 100 is the best score) and quality.

Different analyses are used to assess the PCR primer pairs and sequencing primers. Each analysis results in a penalty based on secondary structures or other potential problems that may have a negative impact on the PCR or Pyrosequencing analysis. The final score of the primer set is a cumulative, weighted sum of the penalties from all analyses.

Within a given template sequence, high scoring primer sets have a higher probability of success during PCR and Pyrosequencing analysis than lower scoring selections. In a high quality primer set, i.e. labeled with color code blue, none of the performed analysis steps have identified any problems of concern. Therefore, a high score primer set can be used directly without any more checks or analyses.

Warnings and penalties

Primers automatically generated by the software are always within the parameters defined in the Default assay settings. However, if the primers deviate too much from the optimal, target settings, penalties and sometimes warnings are issued. Warnings and/or penalties are also issued if the software detects non-favorable conditions, such as secondary structures, in an analysis step. A warning is only issued for potentially serious problems, where the generated penalty is above a certain threshold value (50).

Primers that are added manually, or edited, will receive warnings and/or penalties when parameters outside the Default assay settings are detected, in addition to the warnings and penalties described above.

Warnings are displayed graphically in the information area, and associated with a warning message. Use this information as a quick indication of which analysis results will need to be inspected more closely in the report. In the primer set report, warnings are displayed together with penalties for the individual analysis steps. See **section 7.2.2 - Warning messages** for a list of possible warnings and corresponding descriptions.

Differences in the analysis steps for the different assay types

The analyses carried out by Assay Design Software when designing primers are:

- Individual PCR primer analyses
- PCR primer pair analyses
- Sequencing primer analyses
- Primer set analyses

Some analysis steps differ between the three assay types: genotyping, allele quantification (AQ), and sequence analysis (SQA). The following table gives an overview of the differences in the analyses.

Analysis steps	Genotyping	Allele quantification (AQ)	Sequence analysis (SQA)
PCR primer pair analysis: Amplicon length	Yes	Yes	Not for sequences containing unknown regions.
Primer set analysis: Generation of sequence to analyze	Yes	Yes	No
Sequencing primer analysis: Homopolymers	Yes, with low penalty level.	Yes, with high penalty level.	No
Sequencing primer analysis: A-nucleotide in polymorphism	No	Yes	No

7.2.2 Warning messages

A warning is issued if an analysis step receives a penalty larger than a certain threshold value (50). The warning serves as an indication that the triggering analysis step has detected a potentially serious problem. The following tables give an overview of possible warnings, and the analysis steps they are linked to. Some of the warnings are only generated for manually entered primers, never for automatically generated primers.

PCR primer analysis

Message	Analysis step
Deviation from optimal T_m	Melting temperature analysis.
Self-annealing duplex detected	Duplex formation analysis.
Hairpin loop structure detected	Hairpin loop analysis.
Mispriming site detected	Mispriming analysis within the entered sequence.
Primer with low complementarity	Primer/template complementarity analysis.
Deviation from optimal 3'-end stability	PCR primer 3'-end stability analysis.
T_m for PCR primer is outside settings	Only for manually entered primers. Melting temperature analysis.
Primer length shorter than min	Only for manually entered primers. Primer length analysis.

PCR primer pair analysis

Message	Comment
Amplicon length outside size limit	Only for manually entered primers. Amplicon length analysis.
Deviation from optimal amplicon size	Amplicon length analysis.
Deviation of %GC in PCR primers and/or amplicon	GC content difference analysis.
Deviation of %GC in PCR primers and/or amplicon is outside settings	Only for manually entered primers. GC content difference analysis.
T_m difference out of range	Only for manually entered primers. Melting temperature difference analysis.
Large T_m difference	Melting temperature difference analysis.
Cross-annealing duplex detected	Analysis of duplex formation between the forward and reverse PCR primers.

Sequencing primer analysis

Message	Comment
Low sequencing primer Tm	Melting temperature analysis.
Primer with low complementarity	Complementarity analysis.
Self-annealing duplex detected	Duplex formation analysis.
Hairpin loop structure detected	Hairpin loop analysis.
A homopolymer is detected adjacent to polymorphism	Homopolymer analysis is only performed for genotyping or allele quantification.
Primer length shorter than min	Only for manually entered primers. Primer length analysis.
Position outside settings	Only for manually entered primers. Analysis of the distance between sequencing primer and target region.

Primer set analysis

Message	Comment
Duplex between sequencing primer and biotinylated PCR primer detected	Analysis of duplex formation between sequencing primer and biotinylated PCR primer.
Mispriming site detected for sequencing primer	Mispriming analysis for sequencing primer within the PCR amplicon.
Hairpin loop structure on biotinylated PCR primer detected	Biotinylated PCR primer hairpin analysis.
Loop structure detected on template	Template loop analysis.

7.2.3 PCR primer analyses

The following analyses are performed on the individual PCR primers:

Analysis	Description	Comments
GC content (%)	Calculates the GC content in percent.	
Complementarity	Analyzes the level of complementarity between the PCR primer and its annealing site.	<ul style="list-style-type: none"> Automatically generated primers are always completely complementary. Manually added or edited primers will receive a penalty and a warning for non-complementary sequence motifs. Non-complementarity is penalized more for the 3'-end of the primer than the 5'-end.
Duplex formation	Detects possible PCR primer self-annealing (fwd-fwd and rev-rev) duplex formations.	If potentially serious duplexes are detected, the primer will receive a penalty > 50 and a warning.
Hairpin loops	Detects possible PCR primer hairpin structures.	If potentially serious hairpin loops are detected, the primer will receive a penalty > 50 and a warning.
Melting temperature	Calculates the deviation between the melting temperature of the PCR primers and the optimal melting temperature. By default, the melting temperature algorithm used is the Nearest Neighbor method.	If the deviation is high, the primer will receive a penalty > 50 and a warning.
Mispriming	Detects alternate annealing sites for each PCR primer on the entered sequence (and reverse complementary sequence).	If potentially serious alternate annealing sites are detected, the primer will receive a Penalty > 50 and a warning.
Primer end stability	Calculates the relative stability (Gibbs free energy) difference between the primer 5'- and 3'-ends.	The PCR primer specificity increases if the 5' end is more stable than the 3' end.
Primer length	Calculates the deviation between the actual primer length and the optimal primer length.	If a manually entered primer is shorter than the minimum setting, it will receive a warning. The primer score is not affected.

7.2.4 PCR primer pair analyses

The following analyses are performed on PCR primer pairs:

Analysis	Description	Comments
Amplicon length	Calculates the deviation of the actual amplicon length from the optimal range.	If the amplicon length deviates significantly from the optimal range, the PCR primer pair will receive a penalty > 50 and a warning.
Duplex formation	Detects possible PCR primer cross-annealing (fwd-rev) duplexes.	If potentially serious duplexes are detected, the PCR primer pair will receive a penalty > 50 and a warning.
GC content difference	Calculates the difference in GC content between forward and reverse PCR primer and between the PCR primers and the amplicon.	PCR primers with a GC differences within the optimal range, 0-10%, will be generated without penalty. GC differences higher than 10%, but lower than 30%, will be generated but given a penalty that increases linearly. At penalty > 50, the primer pair will receive a warning.
Melting temperature difference	Calculates the deviation between the actual T_m difference between the forward and reverse PCR primers and the optimal T_m difference range.	If the deviation is large, the PCR primer pair will receive a penalty > 50 and a warning.

7.2.5 Sequencing primer analyses

The following analyses are performed on sequencing primers:

Analysis	Description	Comments
GC content (in %)	Calculates the GC content in percent.	
Complementarity	Analyzes the level of complementarity between the sequencing primer and the annealing site.	<ul style="list-style-type: none"> Automatically generated primers always have complete complementarity. Manually added or edited primers will receive a penalty and a warning for non-complementary sequence motifs. Non-complementarity is penalized more for the 3'-end of the primer than the 5'-end.
Duplex formation	Detects possible sequencing primer self-annealing duplex formations, which can cause background signals in the Pyrosequencing analysis.	Primarily extendable duplexes, i.e. duplexes that are complementary in the 3' end, are penalized. If serious duplexes are detected, the primer will receive a penalty > 50 and a warning.
Hairpin loops	Detects possible sequencing primer hairpin structures that can cause background signals in the Pyrosequencing analysis.	If serious hairpin loops are detected, the primer will receive a penalty > 50 and a warning.
Melting temperature	Calculates the deviation between the melting temperature of the sequencing primer and the optimal melting temperature.	If the T_m is lower than the optimal T_m , the primer will receive a penalty. Penalties > 50 will trigger a warning.
Primer length	Calculates the deviation of the actual primer length and the optimal primer length.	If the primer is shorter than the optimal length, it will receive a warning. The primer score is not affected.
Target distance	Calculates the distance between the sequencing primer and the target region.	
Homopolymers	<p>Homopolymer analysis is only performed for genotyping (medium score weighting) and allele quantification (high score weighting).</p> <p>Detects if the polymorphisms in the target contain adjacent homopolymeric sequences.</p>	<p>Analysis in PSQ software will be more difficult if homopolymeric regions are adjacent to polymorphisms.</p> <p>Sequencing primers annealing over the homopolymeric region are favored because the homopolymeric sequence effect is reduced.</p>

Polymorphism	Polymorphism analysis is only performed for allele quantification (AQ). Checks for A-nucleotide(s) in the target polymorphism.	If an A-nucleotide(s) is/are contained in the target polymorphism, primers that result in the incorporation of T will be favored over primers resulting in incorporation of A.
---------------------	---	--

7.2.6 Primer set analyses

Analysis	Description	Comments
Biotinylated PCR Primer Hairpins	Detects hairpin structures on the biotinylated PCR primer, which may cause background signals in the Pyrosequencing analysis.	If serious hairpin structures are detected, the primer set will receive a penalty > 50 and a warning.
Duplex formation	Detects sequencing primer and biotinylated PCR primer cross-annealing duplex formations, which may cause background signals in the Pyrosequencing analysis.	If serious duplex structures are detected, the primer set will receive a penalty > 50 and a warning.
Mispriming	Detects alternative annealing sites for each sequencing primer within the amplicon.	<ul style="list-style-type: none"> Mispriming will only be detected on the biotinylated strand (because the non-biotinylated strand is removed during the sample preparation phase). If serious alternative annealing sites are detected, the primer set will receive a penalty > 50 and a warning.
Template loops	Detects possible loop structures in the biotinylated strand (the template sequence for Pyrosequencing analysis), which may cause background signals in the Pyrosequencing reaction. An extra A is automatically added to the 3'-end of the amplicon for an additional template loop check, as Taq polymerase frequently adds an extra A to the 3'-end of the amplicon during PCR.	Template loop formation can cause self-priming resulting in background signals in the Pyrosequencing reaction. If serious, extendable template loops are detected, the primer set will receive a penalty > 50 and a warning.
Sequence to analyze	A sequence to analyze is only generated for genotyping and allele quantification assays. A sequence to analyze is generated for the import into PSQ 96MA or PSQ HS 96A SNP software.	The sequence to analyze does not influence the primer set scoring.

7.3 Melting temperature

The melting temperature of a primer depends, among other things, on salt concentration, strand concentration, sequence and length. For PCR primers, one of two algorithms can be selected for the calculation of T_m : **Nearest Neighbor** (default) or **2 × AT + 4 × GC**. For sequencing primers, only the Nearest Neighbor algorithm is used.

7.3.1 Methods for calculating the melting temperature (T_m)

- *Nearest Neighbor*

The melting temperature (T_m) of an oligonucleotide duplex is calculated using the nearest neighbor thermodynamics approach (Rychlik et al., 1990; SantaLucia, 1998) and the following equation:

$$T_m(\text{Primer}) = \frac{\Delta H}{\Delta S + R \times \ln C} + 16.6 \log M - 237.15$$

ΔH	Enthalpy for helix formation.
ΔS	Entropy for helix formation.
R	Molar gas constant (1.987 cal/°C X mol).
C	Concentration of the probe.
M	Molar concentration of monovalent cations.

Values of ΔH and ΔS used (Breslauer et al., 1986) were obtained in 1 M NaCl. The values used for molar concentration of monovalent cations (M) and primer concentration (C) are 150 mM and 0.2 μM , respectively, for PCR primers, and 50 mM and 0.33 μM , respectively, for sequencing primers. The default PCR primer concentration of 0.2 μM can be changed in the **Default assay settings** dialog.

- *2 × AT + 4 × GC*

This equation adds 2 °C for each A and T, and 4 °C for each G and C nucleotide. It is a simple but less accurate method for primer T_m calculation. In this approach the concentration of nucleic acid is not taken into account.

$$T_m = [2 \text{ °C} \times (\text{number of A and T bases})] + [4 \text{ °C} \times (\text{number of G and C bases})]$$

7.4 Guidelines for PCR primer design

When using Assay Design Software to generate primer sets (PCR and sequencing primers), a number of parameters are taken into account. The following section describes some of the underlying knowledge that has been incorporated in the analysis steps of Assay Design Software to select suitable PCR primer sets for PCR and Pyrosequencing analysis.

7.4.1 General guidelines to apply when designing PCR primers

Guidelines	Comments
<p>Primer length</p> <p>The PCR primers should typically be between 18 and 24 bp in length.</p>	<p>The minimum and maximum PCR primer lengths can be changed. However, remember that the shorter the PCR primer, the greater is the risk that it matches more than one region in the genome, thereby increasing the risk of non-specific amplification.</p>
<p>GC content</p> <p>The PCR primers and the PCR product should have approximately the same GC content (in %). The typical GC content of PCR primers ranges from 40% to 60%. For good specificity, the primers should preferably be more GC-rich in the 5'-end and less in the 3'-end.</p>	<p>The allowed difference in GC content can be changed.</p>
<p>Melting temperature</p> <p>The standard range for the melting temperature (T_m) is 60-70 °C. The default settings in Assay Design Software are 56-86 °C for PCR primer T_m. Forward and reverse primer should have similar melting temperatures. By default, the nearest neighbor method is used for calculation of T_m.</p>	<p>Method for T_m calculation can be changed. Furthermore, the minimum and maximum T_m, and the allowed T_m difference between primers, can also be changed.</p>
<p>Amplicon length (PCR product length)</p> <p>Whenever the PCR amplicon size can be directed by primer design, PCR products should be as short as possible, preferably less than 250 bp. The optimal range for PCR amplicon size is 50 to 250 bp. Nevertheless, up to 600 bp long fragments have been tested with good results, and even longer amplicons work for some assays. In general, shorter PCR products have several advantages compared to longer ones, e.g. the amplification efficiency increases and the risk of mispriming or secondary structure formation is reduced.</p>	<p>The optimal amplicon range and the maximum amplicon length allowed can be changed.</p>
<p>Primer dimers/duplexes and internal secondary structures</p> <p>The Assay Design Software checks the PCR primers for dimers and hairpins. Excess biotinylated primer in the PCR reaction can disturb the subsequent Pyrosequencing reaction. There it can cause background if it can form a hairpin loop with a 5' overhang, or a duplex with the sequencing primer.</p> <p>See Chapter 8 - Guidelines for PCR setup and PCR optimization for how to avoid formation of primer dimers/duplexes and internal secondary structures.</p>	<p>Possible dimer and hairpin structures are detected by the software and displayed in the Primer set details and in the report.</p>

Mispriming

To obtain specific amplification, it is important to select PCR primers that do not have alternate annealing sites on the template sequence.

Potential, alternate annealing sites within the entered sequence are detected by the software and displayed on the **Sequence** tab and in the report.

Primer specificity

Primers with a stable 5'-end (high $-\Delta G$ value) and a relatively unstable 3'-end (low $-\Delta G$ value) typically perform best because they are more stable and specific, and thereby less prone to mispriming.

The relative stability is calculated in the Primer end stability analysis step. The ΔG graph on the **Sequence** tab in the **Assay** window visualizes ΔG values for the 5'- and 3'-ends of primers.

7.5 Guidelines for sequencing primer design

A number of parameters are taken into account when using Assay Design Software to generate primer sets (PCR and sequencing primers). The following section describes some of the underlying knowledge that has been incorporated in the analysis steps of Assay Design Software to select suitable sequencing primers for Pyrosequencing analysis.

7.5.1 General guidelines for sequencing primer design

Guidelines	Comments
Primer length The sequencing primer should typically be between 15 and 20 bp, but longer sequencing primers can also be used.	The minimum and maximum sequencing primer lengths can be changed.
Melting temperature Because the Pyrosequencing reaction takes place at 28 °C, the T_m of the sequencing primer can be lower than for the PCR primers. The default target T_m of the sequencing primer is 50 °C. The lowest possible T_m is not absolutely defined, but primers with a calculated T_m of around 40 °C have been used with very good results. However, if using Single Strand Binding protein (SSB) in the Pyrosequencing reaction, a slightly higher primer T_m is required and as low a T_m as 40 °C cannot be recommended for use with SSB. The T_m is calculated using the nearest neighbor method.	
Primer dimers/duplexes and internal secondary structures As the sequencing reaction is run at 28 °C, it is crucial to check the sequencing primer for self-annealing, especially at the 3'-end. <ul style="list-style-type: none"><i>Avoid sequencing primer duplex formation</i> Sequencing primers should be analyzed with regard to their ability to form duplexes. Primers with four or more complementary nucleotides in the 3'-end, and	Possible dimer and hairpin structures are detected by the software and displayed in the Primer set details and in the report.

with 5'-overhang, should not be used. Blunt-ended duplexes will not give rise to background. However, if they have many stabilizing bonds they might self-anneal to a high degree and lower Pyrosequencing signals. Three complementary nucleotides in the 3'-end are acceptable as long as other complementary nucleotides within the primer do not stabilize the duplex.

- *Avoid 3'-end hairpin loops*

Sequencing primers should be analyzed with regard to their ability to form hairpin loops. At 28 °C, as little as three complementary nucleotides in the 3'-end may give rise to background. If hairpin loops cannot be avoided then it may be possible to shorten the primer to give a blunt-end hairpin that cannot generate background signal.

Positioning of the sequencing primer

The sequencing primer should preferably be positioned with its 3'-end as close to the target region as possible, typically within about 5 bp. However, for multiplex design the sequencing primers may have to be moved back further from the target region.

For analysis of single-base In/Dels, the selected primer should preferably be moved back from the target region to generate at least one reference peak before the variable position.

Mispriming

Avoid 3'-end mispriming

Sequencing primers should be analyzed with regard to their ability to misprime within the PCR amplicon. A primer that has six or more complementary nucleotides in the 3'-end at an alternative priming site (and be extra careful with GC-rich 3'-ends) should not be used.

By default, the sequencing primer is positioned between 0 and 3 nucleotides away from the target region. This can be changed.

Potential, alternative annealing sites within the amplicon are detected by the software and displayed on the **Sequence** tab and in the report.

7.5.2 Sequencing primer design for genotyping and allele quantification

Guidelines	Comments
<p>Positioning of sequencing primer</p> <p>The positioning of the sequencing primer is flexible within about 0-15 nucleotides from the polymorphism.</p> <p>Genotyping analysis in PSQ 96MA Software and PSQ HS 96A Software is, in some instances, improved by including a reference peak before the variable position. This is specifically the case when analyzing single-base insertion/deletion polymorphisms.</p> <p>Avoid homopolymers at the polymorphic position</p>	<ul style="list-style-type: none"> • The default setting in Assay Design Software is a primer distance from target of 0-3 bp. • The SNP can be sequenced on either strand.

If the polymorphism is located in a homopolymer, the software will select a primer with a 3'-end that overlaps the homopolymer region. Avoiding homopolymers is especially critical for AQ analysis.

The software will penalize homopolymers harder when assay type AQ has been selected, than when Genotyping has been selected.

For AQ: Avoid A-peaks

Reason for avoiding A-peaks

The use of dATP α S in the Pyrosequencing reaction results in A-peaks that are 10-20% higher than for the other three nucleotides. This must be corrected for by measurement on a heterozygote sample when doing allele frequency measurements on polymorphisms containing A.

Avoidance of A-peaks is only considered when assay type AQ has been selected.

To avoid A-peak height corrections, the software chooses the opposite strand for sequencing primer design whenever possible. However, it is better to have a high quality sequencing primer from which A is read, than a poor primer that generates nonspecific background, but from which T is read.

7.5.3 Sequencing primer design for SQA

Guidelines	Comments
<p>Positioning of sequencing primer</p> <p>The positioning of the sequencing primer should be as close as possible to the sequence to be read. This will maximize read lengths.</p> <p>Nevertheless, it is recommended to start the sequencing with 2-3 known bases, preferably single bases. I.e. include 2-3 known bases flanking the unknown sequence in the target region. These bases, as well as known sequence motifs anywhere along the sequence, can be utilized by the PSQ software algorithm when calling the unknown sequence.</p> <p>Primers for sequence analysis of cloned material</p> <p>Position the sequencing primer in the multiple cloning site. Select the target region so that it includes 2-6 bases before the start of the insert.</p> <p>In this way, the first bases of the sequence are known. Directed dispensations of these bases may improve the sequence quality.</p>	<p>Consider sequencing both strands. In some cases, it may be useful to perform Pyrosequencing reactions in both directions and gather complementary sequence information in order to confirm the sequence.</p>

8 Guidelines for PCR setup and optimization

8.1 Guidelines for PCR setup and optimization for Pyrosequencing analysis

To set up a PCR reaction producing a suitable PCR product for Pyrosequencing analysis, follow the general guidelines below. Optimize the PCR protocol and conditions to obtain good PCR results.

Parameter	Guidelines
PCR primers	<ul style="list-style-type: none"><i>Primer concentrations</i><p>In general, a PCR primer concentration of 0.2 μM is recommended. Lower concentrations may be exhausted before the reaction is completed, resulting in lower yields of the PCR product. However, biotinylated PCR primer concentrations should be kept low to avoid interference in the Pyrosequencing analysis.</p><p>We strongly recommend that the biotinylated primer is purified by HPLC, or an equivalent procedure, to minimize the amount of free biotin.</p><i>Storage of primers</i><p>Biotinylated PCR primers are particularly sensitive to storage. Keep a stock primer solution at $-20\text{ }^{\circ}\text{C}$. For the working solution, prepare small aliquots of diluted primers (10 μM) and store at $-20\text{ }^{\circ}\text{C}$.</p>
PCR product	<ul style="list-style-type: none"><i>PCR product and size</i><p>If possible, select PCR primers to form a PCR product ≤ 250 bp.</p><p>The typical range for PCR amplicon size is 40 to 250 bp. Smaller PCR products have several advantages, e.g. the amplification is easier, and the risk of background in the Pyrosequencing analysis is reduced. However, up to 600 bp long products have been tested with good results.</p><i>GC-rich PCR products</i><p>Amplification of very GC-rich regions ($>70\%$) often benefits from adding 5% dimethylsulfoxide (DMSO) and/or exchanging part of the deoxyguanosine with deoxyinosine. A ratio dI:dG of 3:1 is a good starting point.</p><p>Optimization</p><p>GC-rich templates often need a higher annealing temperature and lower MgCl_2 concentration to amplify well (because high salt concentrations will stabilize secondary structures).</p><i>Checking the PCR product</i><p>On agarose</p><p>Check an aliquot of the PCR product on a 1.5% agarose gel. There should be a clear, strong product band without excess primers, primer-dimers or other non-specific products.</p><p>On PSQ 96MA System</p><p>Use 15-25 μl of the PCR product and 16 pmol of sequencing primer</p>

to give strong signals (single-peak heights of about 15 to 25 units).

On PSQ™HS 96A System

Use 5-10 µl of the PCR product and 3.6 pmol of sequencing primer to give strong signals (single peak heights of ~100-200 units).

DNA

The DNA material should be purified and of high molecular weight (good integrity). The recommended amount of genomic DNA for a standard 25 µl or 50 µl PCR reaction is a minimum of 10 ng of DNA. With smaller amounts than 10 ng DNA in a PCR reaction, there is a risk that there will be too few copies of the genome to give an accurate and robust representation of the genotype/allele content in the sample (resulting in false or skewed genotypes).

MgCl₂ concentration

- *General*

Mg²⁺ ions bind to both nucleotides and DNA, and the concentration of free Mg²⁺ ions therefore depends on the concentrations of compounds like nucleotides, template DNA, free pyrophosphate (PPi) and EDTA (from certain buffers). These compounds bind to the ions via their negative charges. Therefore, the concentration of Mg²⁺ should always be higher than the concentration of these compounds.

- *Mg²⁺ Effect on stringency*

In general, increased Mg²⁺ concentrations will lead to an increased efficiency of the DNA polymerase and therefore to a higher incorporation rate, but may also increase non-specific amplification and reduce fidelity. Excess Mg²⁺ in the reaction can increase non-specific primer binding and increase non-specific incorporation. It may also stabilize secondary structures in the DNA template. This may decrease the amplification efficiency, particularly for GC-rich templates.

Lowered magnesium concentrations will generally make the amplification reaction more stringent, but also less efficient, leading to lower yields. Too little Mg²⁺ in the reaction can result in a lower yield of the desired product.

- *Optimization of Mg²⁺ concentration*

The optimal Mg²⁺ concentration, which may vary from 1 mM to 3 mM, should be determined empirically, while DNA and nucleotide concentrations are kept constant. Perform a magnesium titration from 1 mM to 3 mM in 0.5 mM increments to determine the optimal magnesium concentration.

PCR cycling conditions

The PCR cycling program should be optimized.

- *Optimization of the PCR cycling program*

Hot start DNA polymerases, such as AmpliTaq Gold (Applied Biosystems) and HotStar Taq (Qiagen) are activated gradually during the amplification reaction and therefore require more cycles than a protocol with ordinary DNA polymerases. For best yield, and consumption of all biotinylated PCR primers, which is important for Pyrosequencing, run 45-50 PCR cycles when using a hot start DNA polymerase compared to 35-50 cycles when using ordinary DNA polymerases.

- *Typical PCR program*

A typical PCR program for PCR products of up to about 300 bp, using

a hot start DNA polymerase:

95 °C 5min; 45x(95 °C 15s, T_a °C 30s, 72 °C 15s); 72 °C 5min, 4 °C

The program takes about 1 hour and 45 minutes to run. For PCR products longer than 300 bp, the extension time at 72 °C may need to be increased.

Annealing temperature, T_a

- *Optimal T_a*

The T_m of PCR and sequencing primers is calculated by the Assay Design Software. The optimal annealing temperature (T_a) for the PCR reaction is likely to be between 5-10 °C below the lowest T_m of the pair of primers to be used. Most primers will anneal efficiently in 30 sec or less, unless the T_a is too close to the T_m or they are unusually long.

The typical T_a range is 54-62 °C.

- *T_a effect on annealing and amplification*

Low T_a

One consequence of having a too low T_a is that one or both primers will anneal to sequences other than the true target, since internal single-base mismatches or partial annealing may be tolerated. This can lead to “non-specific” amplification and a consequent reduction in yield of the desired product if the 3'-most base is paired with a target.

High T_a

A consequence of a too high T_a value is that too little product will be amplified, since the likelihood of primer annealing is reduced. Another important consideration is that a pair of primers with very different T_a may never give appreciable yields of a unique product, and may also result in inadvertent “asymmetric” or single-strand amplification of the most efficiently primed product strand.

8.2 PCR setup and optimization - Specific for allele quantification

Application	Guidelines
Allele quantification	<i>Optimization of PCR conditions</i> PCR conditions should be optimized to give a PCR product of high quality, with a yield of at least a 50% and one specific band on an agarose gel. It is very important not to use more primer than necessary in the PCR reaction ($\leq 0.2 \mu\text{M}$). Excess biotinylated primer may result in decreased specific signals and may also give rise to background signal. At least 10 ng genomic DNA should be added to the PCR reaction. This will ensure that enough copies of both alleles are included in the amplification reaction to result in a correct representation of the allele frequency distribution.

8.3 PCR protocol example

As a standard, use 10 ng genomic DNA in a 50 μ l PCR reaction and 125 μ M of each nucleotide. Below is an example of what a typical PCR reaction mix can look like. The example shows a PCR reaction mix (2.0 mM MgCl₂) using AmpliTaq Gold, for one and ten 50 μ l reactions respectively. The volumes are in microliters. Use 45 μ l of reaction mix and 5 μ l 2 ng/ μ l genomic DNA per tube/well.

PCR mix component	Volume per reaction (μ l)		End conc./amount in a 50 μ l PCR reaction
	1 tube	10 tubes	
H ₂ O	31.2	312	-
10x PCR buffer II (Applied Biosystems)	5	50	1x
MgCl ₂ (25 mM)	4	40	2.0 mM
dNTP (2.5 mM)	2.5	25	0.125 mM each
Forward PCR primer (10 μ M)	1	10	10 pmol
Reverse PCR primer (10 μ M)	1	10	10 pmol
AmpliTaq Gold® (Applied Biosystems)	0.3	3	1 U
Total:	45	450	-
Template DNA (2 ng/ μ l)	5	5 each	10 ng

8.3.1 Optimization of the PCR protocol and conditions

For optimization, start with the parameters annealing temperature (T_a) and MgCl₂ concentration.

The annealing temperature typically falls in the range 54 $^{\circ}$ C – 64 $^{\circ}$ C, and the MgCl₂ concentration is typically between 1 mM and 3 mM. Select two DNA samples that can be used for all PCR optimizations. A good starting point is to try three different temperatures (e.g. 54 $^{\circ}$ C, 57 $^{\circ}$ C, and 60 $^{\circ}$ C) and three different MgCl₂ concentrations (1.5 mM, 2.0 mM and 2.5 mM) while keeping all other parameters constant.

9 Hints & Tips

9.1 Tips for succeeding with difficult assays

The software performs a number of analyses to differentiate between primer set candidates and thereby identify the best possible assay to use. However, for some assays it may be difficult to find an assay without flaws. It is then often possible to work around, or resolve, the detected, potential problem by using one of the tips below.

For information on which analyses the software performs, as well as guidelines on PCR and sequencing primer design, see **Chapter 7 - Guidelines for primer design**. For guidelines on PCR setup and optimization, see **Chapter 8 - Guidelines for PCR setup and Optimization**.

Issue	Tips
Non-specific PCR amplification	<p><i>Perform a homology search of the DNA sequence</i></p> <p>Avoid placing PCR primers in highly homologous sequence regions. For DNA analysis of genes in homologous sequence regions, it may be wise to use a search engine (e.g. BLAST, Basic Local Alignment Search Tool) to check possible homologous regions such as pseudogenes in the given genome. Generated PCR primers can also be checked for sequence homology. One such homology search engine can be found on the website of the National Center for Biotechnology Information (NCBI) at: http://www.ncbi.nlm.nih.gov/BLAST.</p> <p>If homologous sequences are found</p> <p>If highly homologous sequences were found during the homology search, the sequences may be aligned using an alignment tool. One such alignment tool can be found on the web site at the web site of Institut National de la Recherche Agronomique (INRA) http://prodes.toulouse.inra.fr/multalin/multalin.html</p>
Template loop formation (self-priming)	<p><i>Add an extra nucleotide to the non-biotinylated PCR primer</i></p> <p>The 3'-end of the template, which can form a loop, is defined by the 5'-end of the non-biotinylated PCR primer. The 3'-end complementarities of a template loop may be removed by modifying the non-biotinylated PCR primer. Enter a random, extra nucleotide to the 5'-end of the non-biotinylated PCR primer in the Primer set area. The score and quality of the primer set will automatically be updated as soon as you leave the field (the new score is shown in the upper right corner of the Primer Set area). Continue to try different nucleotide-additions until the loop has disappeared and the score has improved.</p>
Hairpin loops, 3'-end duplexes, or misprimings that cannot be avoided	<p><i>Check if generated background signal would affect the polymorphic position</i></p> <p>If hairpin loops, 3'-end duplexes, or sequencing primer misprimings cannot be avoided, check what nucleotides will get incorporated to determine if the generated</p>

Hairpin loops or 3'-end duplexes that cannot be avoided

background signal would affect the target region to be analyzed. Also check if the dispensation order could be modified to avoid the background signal from showing up in the variable positions.

Modify the sample preparation procedure

If the sequencing primer forms hairpins or duplexes, add an extra wash step to the sample preparation to remove background signal. After annealing, transfer the beads to a new PSQ 96 MA or PSQ HS 96A Plate containing 40 µl or 10 µl respectively of fresh 1x Annealing buffer per well, leaving excess sequencing primer behind.

9.2 Tips for avoiding PCR cross-contamination

- Set up physically separated working places for template preparation, PCR set-up, and post PCR analysis.
- Use dedicated (PCR use only) pipettes, micro-centrifuges and disposable gloves.
- Use aerosol resistant pipette tips.
- Set up a PCR reaction under a laminar flow hood equipped with UV light.
- Use sterile techniques and always wear fresh gloves.
- Always use new and/or sterilized glassware and plastics to prepare the PCR reagents and genomic DNA.
- Use PCR reagents and solutions only for PCR reactions, and store these reagents in small aliquots.
- Always include a negative control (all reaction components except DNA) and a positive control (e.g. a sample that has been successfully amplified in previous experiments).

9.3 Tips for assay controls

Parameter	Tips
Controls in the PCR reaction	<ul style="list-style-type: none">• Always include a negative control that includes all reaction components except DNA.• When setting up a new assay, perform PCR optimization on a couple of control DNA samples that have been successfully amplified in previous experiments.
Controls in the Pyrosequencing reaction	<ul style="list-style-type: none">• Sequencing primer only.• DNA template only (without sequencing primer)• PCR negative control (with sequencing primer)

9.4 Tips for multiplex assay design

The Assay Design Software does not support automatic multiplex assay design. However, it contains some useful support functions for multiplex design.

- Duplex design: Once a candidate primer set has been selected for one polymorphism, check for mispriming of the primers in the template containing the second polymorphism, and vice versa. Do as follows:
 - Open two assay windows and enter the two sequences of interest.
 - Design primer sets for assay one.
 - Select the desired primer set in the **Primer set** list, right-click and select **Copy primer set** from the right-click menu.
 - Switch over to the second assay window. Select biotinylation of the same primer as for assay one. This can for example be done by clicking the button **Set sequencing primer direction** in the **Primer set** list.
 - Put the cursor in the top field of the **Primer set** list, i.e. in the forward PCR primer field, right-click and select **Paste** from the right-click menu. The selected primer set from assay one will now be pasted into the fields of assay two.
 - As soon as you leave the entry fields, the primer set score is updated. The score can be expected to be zero (discard), since the primers from assay one are not expected to be complementary to the sequence of assay two. Look in the info field on the **Sequence** tab to see if the PCR primers from assay one form any serious misprimings in the sequence of assay two.
- When designing the different primer sets in Assay Design Software, increase the “Maximum Distance from Target” in the settings in order to obtain sequencing primers at varying 3'-positions. This will allow for a higher flexibility in design of the multiplex assay. Another option is to widen the target region, so that it includes a number of nucleotides preceding the polymorphism. The program is thereby forced to generate sequencing primers further away from the SNP.

For more information about multiplex assay design, visit <http://techsupport.pyrosequencing.com>.

9.5 Tips for universal biotinylated PCR primers

It is possible to design and analyze PCR primers with a universal tail for use with a universal biotinylated PCR primer. This allows the use of the same biotinylated primer in different PCR reactions/assays.

To design PCR primers with a universal primer tail:

- Generate primer sets (or enter a previously designed primer set) according to the description in **Chapter 4**.
- Select a primer set by clicking in the **Primer set** list.
- Position the cursor in the biotinylated primer field, at the top of the **Primer set** area. Edit the biotinylated PCR primer to include a certain number of additional nucleotides at the 5'-end, which make up the universal tail. This can be done either by Copy/Paste or by typing in the appropriate sequence.

- As soon as you leave the primer entry field, the score and quality will automatically be updated. The new score and quality are displayed at the top right corner of the **Primer set** area. The **Sequence** tab of the **Assay window** has been updated to display the primer with its associated tail.

Note: A primer set with a universal tail can be expected to receive low score and quality because the tailed primer will receive a substantial penalty for having a high T_m value as well as for having low complementarity to the DNA template sequence. This is a natural consequence of the addition of a long tail, and does not mean that the primer is unsuitable for Pyrosequencing analysis. The score and quality values should therefore be ignored in this case, since they are based on partially irrelevant analyses. Instead, the important thing is to check if any serious new misprimings, PCR duplexes or template loops have been formed and detected.

9.6 Tips for using the Gibbs free energy (ΔG) graph

The Gibbs free energy is used to determine a primer's priming specificity. The ΔG values indicate how specific the primer is and how efficiently it will anneal to its intended target. Primers with a stable 5' end i.e. low ΔG value and a relatively unstable 3' end, i.e. high ΔG value, typically perform best because they are both more stable and specific and thereby rarely misprime. A primer with low stability at its 3'-end will function well in a PCR because the base pairing close to the 3'-end with non-target sites are not sufficiently stable to initiate extension (false priming). Conversely, primers with a stable 3'-terminus need not anneal with the target along their entire length in order to prime efficiently. This could result in non-specific products (false priming).

Assay Design Software assesses the relative stability of the PCR primers in the "Primer end stability" analysis. For this, the following software settings are used:

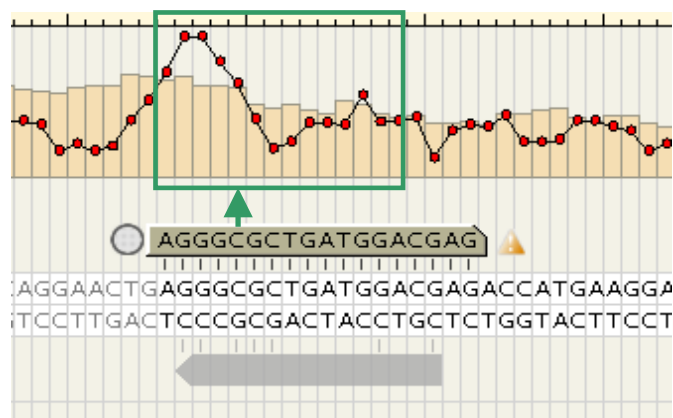
Optimal ΔG difference: 2.0

Maximum ΔG at 5'-end: -8.0

Maximum ΔG at 3'-end: -5.0

Minimum ΔG at 3'-end: -11.0

In addition, the ΔG graph can be used to visualize the ΔG distribution in the different primers. Here, each data point represents the stability of a nucleotide pentamer. The data points correspond to the Gibbs free energy of the nucleotide in that position and the four nucleotides immediately downstream in the sequence.



The Delta G graph of the selected primer will include graph points from the 5'-end of the primer to 5 nucleotides upstream from the 3'-end (boxed), as each point represents the Delta G of a nucleotide pentamer.

9.7 Tips for using the melting temperature (T_m) graph

Each bar in the T_m graph represents the T_m for a primer of 20 bp by default with its 5'-end in that position.

To view T_m for a selected primer of any length, click on the wanted primer in the graphical representation on the Sequence tab. This automatically adjusts the T_m graph to this primer length. The T_m of this particular primer is represented by the bar at the 5'-position.

The primer T_m can be adjusted by changing the primer length or the primer position. This can be done either by adding/removing nucleotides in the **Primer set** area, or by dragging and dropping or resizing the primer on the **Sequence** tab.

9.8 Tips for analyzing InDels in homopolymeric stretches

If the polymorphic nucleotide(s) of an InDel can form a homopolymer with adjacent nucleotides (i.e. three or more of the same nucleotide in a row), then a sequencing primer should be selected that overlaps the homopolymeric nucleotides. Enter two different notations for forward or reverse assays (see Example 1).

Example 1

Allele 1: TTTTT (insertion)
Allele 2: TTT (deletion)

Forward assay

Enter sequence as:

TTT[TT]

Select [TT] as target region. Open Settings for the current assay and, for the sequencing primer, check only the "Generate forward primers" box.

Reverse assay:

Enter sequence as:

[TT]TTT

Select [TT] as target region. Open Settings for the current assay and, for the sequencing primer, check only the "Generate forward primers" box.

9.9 Analyzing short tandem repeats (STRs)

To design an assay for STR analysis, use the InDel notation (see Example 2).

Example 2:

Allele 1: ACGACGACG

Allele 2: ACGACGACGACG

Allele 3: ACGACGACGACGACGACG

Assay Design Software will not position primers inside the selected target region. By default, sequencing primers will be positioned 0 to 3 nucleotides from the border of the selected target area. In order for the sequencing primer to overlap some part of the constant repeat region, two different notations and choices of target region should be used for forward and reverse sequencing primer generation respectively.

Forward assay:

ACGACGACG[ACG][ACGACG]


The target region should be selected starting four nucleotides into the constant part of the repeat region. In the example above, choose CGACG[ACG][ACGACG] as target region and generate only forward sequencing primers (Tick the box "Generate forward primers" in the Current assay settings dialog). By choosing this target area, the forward sequencing primers that are generated will overlap between 0 and 4 bases of the constant repeat region.

Reverse assay:

[ACGACG][ACG]ACGACGACG

The target region should be selected starting four nucleotides into the constant part of the repeat region. In the example above, choose [ACGACG][ACG]ACGAC as target region and generate only reverse sequencing primers (Tick the box "Generate reverse primers" in the Current assay settings dialog). By choosing this target area, the reverse sequencing primers that are generated will overlap between 0 and 4 nucleotides of the constant repeat region.

9.10 Troubleshooting guide

Problem	Possible cause	Solution
Generated primers are not shown in (or disappear from) the sequence graph of the Sequence tab.	Inadvertent use of the buttons to show/hide primers in the sequence graph.	Click the appropriate button to the left of the sequence graph to show the primer(s) that is/are hidden.
Primers are positioned incorrectly in the sequence graph of the Sequence tab.	The 3'-end of an edited primer is positioned at the same site as the 3'-end of the original primer, even if nucleotides have been added or changed at its 3'-end. If the contents of an entry field in the Primer set area are changed by deleting the primer sequence and then typing/pasting a new sequence in the field, the software considers this as an edited primer. Thus, it does not reposition the primer on the Sequence tab.	To enter new primers in the Primer Set area, press the New Primer Set button  to blank the primer entry fields before the new sequence string is entered/pasted.
The application becomes slower and slower	The application has been running for a long time without restart, which may lead to consumed memory resources.	Restart the application to relieve memory.
The same PCR primer mispriming or template loop is shown twice in the report.	If the sequence contains multiple polymorphisms, each haplotype is analyzed separately. I.e. if a PCR primer misprimers over an SNP, the mispriming to both alleles will be shown in the report, thereby seemingly showing the same mispriming twice.	No action is needed.
No PCR primers can be generated.	The sequence entered on the Sequence Editor tab may be too short to allow positioning of suitable PCR primers.	Enter a longer sequence and make sure that the target polymorphism/sequence is flanked by a sufficiently long sequence stretch. 200 bp on either side of the target is recommended.
Import in PSQ software fails because an Entry with the same name already exists in the database.	Changing the name of the .xml file in Windows Explorer before import in PSQ software will not solve the problem. The Entry name is stored inside the .xml file.	Open the assay file in Assay Design Software. Choose File Save As and save the assay with a new name. The new assay name will be stored inside the .xml file. Retry import in PSQ software.

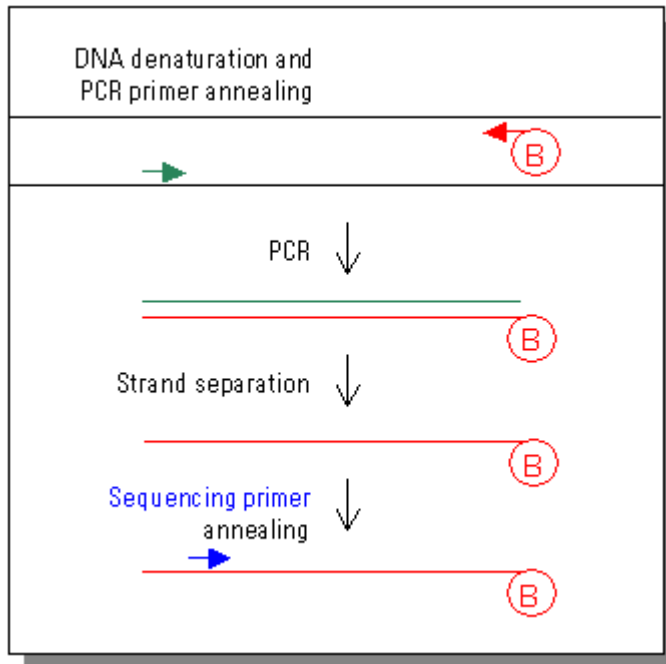
<p>Batch analysis fails for one or several assays.</p>	<p>If batch analysis fails with one of the three following messages the problem is caused by illegal characters within the sequence, or lack of a target polymorphism. Messages:</p>	<p>Confirm that all assays contain polymorphisms (or SQA target regions). Resolve any incorrect characters in the sequence and re-start batch analysis.</p>
	<p><i>Skipped empty or un-intelligible assay 'test_48'.</i></p>	
	<p><i>'Test_48': Before running the analysis, target area or polymorphism has to be defined.</i></p>	
	<p><i>Before running batch resolve the following errors: Something went wrong when interpreting the sequence.</i></p>	
<p>A high scoring assay could not be run in PSQ 96MA or PSQ HS 96A.</p>	<p>Example: an AQ assay of three consecutive SNPs that share allelic bases (e.g. ...MRS...) is not possible to analyze in PSQ because of confounded signals. This is not flagged until the assay is imported/entered in PSQ system software, because Assay Design Software does not have the ability to generate dispensation orders.</p>	<p>Enter/import assays in PSQ system software, generate dispensation orders and confirm that assays are analyzable, before primers are ordered.</p>

10 Appendix A. Methodological background

10.1 Sample preparation

The starting material for a Pyrosequencing™ reaction is a PCR-amplified, single-stranded DNA template with a sequencing primer hybridized to it. Several methods to generate templates for Pyrosequencing analysis have been described (e.g. Ronaghi et al., 1996; Nordström et al., 2000, Nordström et al., 2002, Diggle and Clarke, 2003).

The principle for PCR and sample preparation prior to Pyrosequencing analysis is outlined below:



When using streptavidin-coated magnetic- or Sepharose™ beads for sample preparation (Ronaghi et al., 1996), one of the PCR primers should be biotin labeled (B) for immobilization to the beads. The other PCR primer should be unlabeled.

As free biotin will compete with the biotinylated PCR product for binding to streptavidin, thereby lowering the signal level, we strongly recommend purifying the biotinylated PCR primer by HPLC, or equivalent procedure, to minimize the amount of free biotin and maximize the proportion of biotinylated primer.

After immobilization, using NaOH denaturation, and annealing of the sequencing primer, the immobilized strand can be analyzed using the Pyrosequencing™ technology.

References:

1. Ronaghi, M., Karamohamed, S., Peterson, B., Uhlén, M., Nyrén, P. (1996). Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry* 242: 84-89.
2. Nordström, T., Nourizad, K., Ronaghi, M., Nyrén, P. (2000). Method enabling Pyrosequencing on double-stranded DNA. *Analytical Biochemistry* 282: 186-193.
3. Nordström, T., Alderborn, A., Nyrén, P.J. (2002). Method for one-step preparation of double-stranded DNA template applicable for use with Pyrosequencing technology. *Biochem Biophys Methods* 31; 52(2): 71-82.
4. Clarke S.C and Diggle, M.A: (2003). A novel method for the preparation of single stranded DNA for Pyrosequencing. *Molecular Biotechnology* 24: 221-224.

10.2 Pyrosequencing systems

10.2.1 Introduction

Two different systems are available for Pyrosequencing analysis:

- **PSQ 96MA System**

This system supports genotyping, allele quantification (AQ) and sequence analysis (SQA).

- **PSQ HS 96A System**

This system is a highly sensitive system that requires less reagents than PSQ 96MA. It supports genotyping and allele quantification (AQ).

The two systems are used to perform the Pyrosequencing reaction on the PCR product. Use Assay Design Software to design an assay suitable for Pyrosequencing analysis. The assays can be saved as .xml files, which can be imported by PSQ 96MA Software version 2.1 and higher and PSQ HS 96A Software version 1.2 and higher.

10.2.2 Genotyping and mutation analysis

The PSQ 96MA Instrument with PSQ 96MA SNP Software or PSQ HS 96A Instrument with PSQ HS 96A Software and SNP reagent kit perform highly accurate and reproducible Single nucleotide polymorphism (SNP) and mutation analysis.

Automatic multiplex genotyping of polymorphisms such as point mutations, insertions/deletions (InDels) and SNPs reduce the cost per accurate result and increase sample throughput.

Multiple SNPs from a short stretch of DNA and di-, tri- and tetra-allelic polymorphisms can be analyzed. The internal control capabilities provide the sequence context around the SNP or mutation.

For more information, see the user documentation for PSQ 96 MA System or PSQ HS 96A System.

10.2.3 Allele quantification (AQ)

Once polymorphisms are identified, the frequency of alleles can be determined using the dedicated module (AQ) within the PSQ 96MA SNP Software or PSQ HS 96A Software. Alleles from both pooled genomic and single mixed population samples can be quantified (calculating allele frequencies in pooled sample populations can increase the efficiency and decrease the cost for large population studies because only a single PCR amplification reaction followed by one sequencing reaction is needed when analyzing the samples).

Allele quantification determination of multiple SNPs, tri- and tetra allelic SNPs and InDels can all be achieved. In multiplex assays, it is possible to obtain the relative peak heights and then calculate the allele frequencies manually.

The linear relationship between peak height and nucleotide incorporations, in combination with high signal-to-noise ratios for the obtained peaks, facilitate the estimation of allele frequencies, even for low frequency SNP alleles.

For more information, see the user documentation for PSQ 96 MA System or PSQ HS 96A System.

10.2.4 Sequence analysis (SQA)

PSQ 96MA Instrument is used together with dedicated software and kits for the analysis of DNA sequences typically 30-50 bp in length. The PSQ 96MA SQA software automatically completes base-calling and the sequence alignment function enables easy comparison to a master sequence. Sequence data can also be exported to other systems for database comparisons.

Along with sequencing short-strand DNA templates, PSQ 96MA System SQA is capable of straightforward sequencing of cloned DNA material, templates with strong secondary structures and complementary DNA. This is achieved by the addition of single-stranded DNA binding protein (SSB) after the primer-annealing step.

To sequence DNA templates, a suitable dispensation order must be entered in the Define SQA Entry dialog in PSQ 96MA SQA software.

For more information, see the user documentation for PSQ 96MA System.

10.3 Definitions

10.3.1 Alleles

An allele is any one of a number of alternative forms of the same gene occupying a given locus (position) on a chromosome.

If an individual is homozygous for a certain polymorphic position, the allele sequences in that position are identical (i.e. the same nucleotide occupies the position). If an individual is heterozygous for a certain polymorphic position, the allele sequences differ at that position.

Example (3 polymorphisms in one gene, reading forward)		
5'GT G GAGAT.....GCTAGCTAA.....TAC C G...3'		Chromosome 1
3'CACCTCTA.....CGATCGATT.....AT G GC...5'- Biotinylated		
5'GT A GAGAT.....GCTAGCTAA.....TAT C G...3'		Chromosome 2
3'CATCTCTA.....CGATCGATT.....AT A GC...5'- Biotinylated		
Heterozygous Genotype: G/A	Homozygous Genotype: T/T	Heterozygous Genotype: C/T

DNA is obtained from both alleles when amplifying the samples with PCR, and will be analyzed simultaneously in the Pyrosequencing reaction.

If there are three possible polymorphic positions (reading forward), for example G/A, T/C and C/T, 27 different theoretical outcomes are possible.

1.	G/G, T/T, C/C	10.	A/A, T/T, C/C	19.	G/A, T/T, C/C
2.	G/G, T/T, T/T	11.	A/A, T/T, T/T	20.	G/A, T/T, T/T
3.	G/G, T/T, C/T	12.	A/A, T/T, C/T	21.	G/A, T/T, C/T
4.	G/G, C/C, C/C	13.	A/A, C/C, C/C	22.	G/A, C/C, C/C
5.	G/G, C/C, T/T	14.	A/A, C/C, T/T	23.	G/A, C/C, T/T
6.	G/G, C/C, C/T	15.	A/A, C/C, C/T	24.	G/A, C/C, C/T
7.	G/G, C/T, C/C	16.	A/A, C/T, C/C	25.	G/A, C/T, C/C
8.	G/G, C/T, T/T	17.	A/A, C/T, T/T	26.	G/A, C/T, T/T
9.	G/G, C/T, C/T	18.	A/A, C/T, C/T	27.	G/A, C/T, C/T

10.3.2 Single nucleotide polymorphisms (SNPs)

Single nucleotide polymorphisms (SNPs) are DNA sequence variations that occur when a single nucleotide (A,T,C or G) in the genome sequence is altered. For example, a SNP might change the DNA sequence AAGGCTAA to ATGGCTAA. Both forms must occur with a frequency of 1% (0.01) or greater in a large population to be classified as a SNP (otherwise it is a random mutation).

SNPs are evolutionarily stable, not changing much from generation to generation, making them suitable to use as genetic markers in population studies.

Two out of three SNPs involve the replacement of cytosine (C) with thymine (T). SNPs occur on average every 100 to 300 bp along the 3-billion-base human genome. SNPs can occur in both coding and non-coding regions of the genome. Many SNPs have no effect on cell function, whereas others could predispose people to disease or influence their response to a drug.

10.3.3 Insertions and deletions (InDels)

An insertion or deletion is a mutation where one or several nucleotides have been inserted/deleted in a DNA sequence. The size of the InDel can vary from single-base to a part of a chromosome.

Homozygous single-base deletion:		Homozygous single-base insertion:	
Allele 1	CA TCCGGA...	Allele 1	CAATCCGGA...
Allele 2	CA TCCGGA...	Allele 2	CAATCCGGA...
Heterozygous single-base deletion:		Heterozygous single-base insertion:	
Allele 1	CA TCCGGA...	Allele 1	CA TCCGGA...
Allele 2	CAATCCGGA...	Allele 2	CAATCCGGA...

InDels in protein coding parts of genes have great impact on the protein's function, activity and structure because they can cause:

- frameshifts (if the InDel is **not** a multiple of three nucleotides)
- addition/loss of one or several amino acids (if the InDel is a multiple of three nucleotides)

Frameshifts

Frameshifts usually result in a non-functioning or malfunctioning protein since the frameshift changes the reading frame (three nucleotides give one amino acid) and thereby alters the protein, which may lose all activity.

Addition/loss of one or several amino acids

When the addition/loss of one or several amino acids occurs, the amino acid chain that builds up the protein changes. This may alter the structure and/or function of the protein.

Homozygous versus heterozygous InDels

If the InDel is heterozygous (see above), there will still be one allele left expressing the correct protein and this can be sufficient in some cases.

If the InDel is homozygous (see above), there will not be any correct proteins expressed, and the protein's function is lost or altered.

10.3.4 Short tandem repeats (STRs)

A short tandem repeat (STR) is a repeated DNA sequence in which the repeat elements are typically two to five base pairs long. Different alleles have different numbers of the repeat element.

10.3.5 Sequence database files

Two of the major nucleic acid sequence databases are:

- GenBank (<http://www.ncbi.nlm.nih.gov/Entrez/index.html>)

and

- EMBL Nucleotide Sequence Database (<http://www.ebi.ac.uk/embl>).

GenBank is an annotated collection of all publicly available DNA sequences and the genetic sequence database of the US National Institute of Health.

The EMBL Nucleotide Sequence Database is Europe's most comprehensive nucleotide sequence database.

Both databases use their own sequence file format, consisting of a header, which contains general information, such as keywords, author names, source, organism etc., and the actual nucleic acid sequence.

Both GenBank and EMBL sequences can be saved in FASTA format (see below).

GenBank sequence format

The Gen-Bank sequence format is identified by the entry **ORIGIN** at the beginning of the nucleic acid sequence and ends with **//**. See **section 4.2.2.2** for detailed information.

EMBL sequence format

The EMBL sequence format is identified by the entry **SQ** at the beginning of the nucleic acid sequence and ends with **//**. See **section 4.2.2.2** for detailed information.

FASTA sequence format

The sequence file format FASTA is used by many sequence alignment and homology search programs. A sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The description line is distinguished from the sequence data by a greater-than ">" symbol in the first column. See **section 4.2.2.2** for detailed information.

10.3.6 Sequence to analyze (genotyping and allele quantification)

The **Sequence to analyze** is a short part of a DNA sequence (complementary to the sequence of the biotinylated strand) that contains one or several polymorphisms to be analyzed using PSQ 96MA System or PSQ HS 96A System. The **Sequence to analyze** always starts with the first nucleotide after the sequencing primer, i.e. where the sequencing reaction starts.

Example of a Sequence to analyze:

AT/CCGTGT[T]CCCA

T/C denotes a di-allelic SNP

[T] denotes an insertion/deletion polymorphism

This is the sequence to analyze for the following DNA sequence with annealed sequencing primer:

5'GGCACGAATCGACTT 3'

3'CCGTGCTTAGCTGAATA/GGCACA [A] GGGT . . . 5'- Biotin

Assay Design Software automatically generates a **Sequence to analyze** for genotyping and allele quantification assays. When importing the assay into PSQ 96MA System or PSQ HS 96A System, a **dispensation order** is generated based on the **Sequence to analyze**, and an SNP Entry is thereafter automatically created containing the **Sequence to analyze** and **dispensation order**.

10.3.7 Dispensation order

The dispensation order determines in which order the nucleotides will be added (dispensed) to the Pyrosequencing reaction by the PSQ 96MA Instrument or PSQ HS 96A Instrument. The **Sequence to analyze**, which has automatically been generated by Assay Design Software for genotyping and allele quantification assays, is used to generate a **Dispensation order** when the assay is imported into PSQ 96MA Software or PSQ HS 96A Software.

Note: The **Dispensation order** for SQA assays must be entered into the PSQ 96MA Software by the user.

Example:

Sequence to analyze: AT/CCGTCAAAGC

will result in the **Dispensation order:** GATCAGTCA

The first nucleotide in the **Dispensation order** as well as the first nucleotide following a polymorphic position is usually a blank internal negative control and is not the same as the first nucleotide in the **Sequence to analyze**.

10.3.8 Reference peaks and quality control window

Reference peaks are used by PSQ 96MA Software and PSQ HS 96A Software to determine which peak height corresponds to the incorporation of a single nucleotide. The single peak height level is then used to determine the multiplicity of all peaks generated in the Pyrogram, and thereby the sequence and genotype of the sample. All non-variable peaks generated in the sequencing reaction with a multiplicity of three or lower are used as reference peaks by the algorithm.

To assess the quality of the sequence, all dispensations within the **quality control window** are used, including negative control dispensations. Reference peaks included in the quality control window will affect the analysis. The quality window will be positioned symmetrically around the polymorphic position. The quality control window is static, so if dispensations fewer than half the value for the quality control window size are available before the SNP, the size of the actual quality window will be reduced. In the case of multiple SNPs analyzed with one sequencing primer, the quality control window for one SNP position will end at the next SNP.

To make the best use of the reference peaks in the Pyrosequencing analysis by PSQ 96MA Software and PSQ HS96A Software, design the sequencing primer to obtain:

- reference peaks as close as possible to the polymorphism
- single peaks as reference peaks
- reference peaks both before and after the polymorphic position, especially when analyzing insertions/deletions or multiple SNPs

10.3.9 Simplex and multiplex assays

With the PSQ 96MA and PSQ HS96A systems it is possible to create two different types of assays in SNP genotyping: simplex and multiplex assays.

- In the **simplex assay** one sequencing primer is used, which means that the polymorphic positions must be in the same sequence and that the first and last polymorphisms should preferentially be within 15 nucleotides (otherwise consider running a multiplex assay). One sequencing reaction is performed in the simplex assay.
- In **multiplex assays**, one sequencing primer per polymorphic position is used in the sequence reaction. The polymorphic positions do not have to be in the same sequence. Several different sequencing reactions occur simultaneously.

Simplex assay

In simplex assays, one or several polymorphic positions are genotyped using one sequencing primer.

PSQ 96MA System and PSQ HS 96A System are optimized for genotyping within 15 nucleotides from the sequencing primer.

Multiplex assay

In multiplex assays, several polymorphic positions are analyzed in the same reaction. The polymorphisms can be positioned in the same DNA fragment or in different DNA fragments. One sequencing primer per polymorphism is used in the sequencing reaction.

PSQ 96MA SNP Software and PSQ HS 96A Software supports duplex and triplex sequencing of single polymorphic positions (di-, tri- and tetra-allelic SNPs, and InDels).

Example 1: Multiplex assay where the polymorphisms are positioned in the same DNA fragment

3' . . . AACATTCGGCTTACAG/ACAGCTGACCTAGCCTCGGATGAAC/TACTTCGTCGAAC . . .
5'TAAGCCGAATGT → 5'TGGATCGGAGCCT →

One DNA fragment with two polymorphic positions that are sequenced at the same time. The peaks come from two different sequencing reactions when the primers are extended.

Example 2: Multiplex assay where polymorphisms are positioned in two different DNA fragments

3' . . . ACGTGGATTACATTCGGCTTACAG/ACAGCTGAGTAAAGTTAGT . . .
5'TAAGCCGAATGT →

3' . . . GTTGACAGGACGTACGCTTACACCGTAAGGTGTGAAC/TACC . . .
5'ATGTGGCATTCCAC →

Two different DNA fragments are sequenced at the same time. The peaks come from two different sequence reactions when the primers are extended.

Note: Assay Design Software does not support the automatic primer design for multiplex assays. To use Assay Design Software for multiplex assay design, design the primer sets separately, and manually check the different primer combinations, e.g. for duplex formation. See **Chapter 9 - Hints & Tips**, for more information.

10.3.10 Mispriming

To obtain specific amplification during PCR, it is important to select PCR primers that do not have alternative priming sites on the DNA template sequence.

Mispriming in the Pyrosequencing reaction occurs when the sequencing primer binds to the wrong part of the sequence, which may result in an overlaid sequence in the pyrogram.

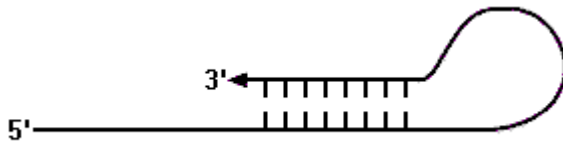
10.3.11 Secondary structures

Double or overlaid sequences, broad peaks or signal compression in the pyrogram may indicate the occurrence of secondary structures in the DNA template.

Secondary structures can hinder the annealing of the sequencing primer, which can result in low and/or spurious peaks.

- *Secondary structures can be the result of self-priming.*

Self-priming occurs when a part of the DNA sequence at the 3' end is complementary to another part of the DNA sequence and the sequence thus binds to itself and acts as a primer.



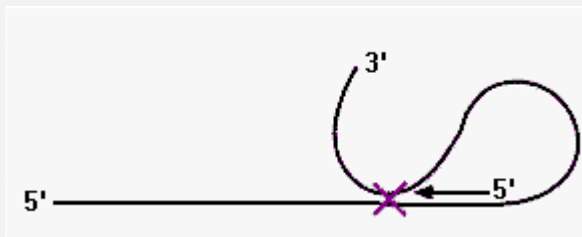
This may result in an overlaid sequence in the pyrogram.

- *Secondary structures in a DNA template can cause problems in the sequencing reaction due to interference with enzymes in the assay.*

Secondary structures may interfere with the enzymes in the assay. Secondary structures can hinder the DNA polymerase from binding to its priming site or obstruct the procession of the DNA polymerase, which could lead to broad peaks and/or incorrect sequencing context in the Pyrogram.

Example

A secondary structure has been formed that inhibits the enzyme's activity.



11 Appendix B. Assay types

11.1 Genotyping and allele quantification (AQ)

11.1.1 Introduction

Assay Design Software can be used to design assays (PCR and sequencing primers) for genotyping, mutation analysis, and allele quantification (AQ) in PSQ 96MA System or PSQ HS 96A System.

This part of the methodology provides information on:

- Polymorphisms for which assays can be designed in Assay Design Software
- How to enter the polymorphism in the Assay Design Software

11.1.2 Polymorphisms for which assays can be designed

Assay Design Software supports assay design for di-, tri- and tetra-allelic single nucleotide polymorphisms (SNPs), point mutations, and insertions/deletions (InDels).

Di-allelic SNPs

A SNP position in a gene is said to be di-allelic if two possible alleles (nucleotides) can occupy the SNP position, for example C or T (denoted C/T or Y when entered in Assay Design software).

Tri-allelic SNPs

A SNP position in a gene is said to be tri-allelic if three possible alleles (nucleotides) can occupy the SNP position, for example G, T or C (denoted G/T/C or B when entered in Assay Design software). A diploid individual can have any combination of two of these nucleotides.

Tetra-allelic SNPs

A SNP position in a gene is said to be tetra-allelic if four possible alleles (nucleotides) can occupy the SNP position, i.e. C, T, G and A (denoted C/T/G/A or N when entered in Assay Design software). A diploid individual can have any combination of two of these nucleotides.

Mutations

A point mutation is a DNA sequence variation that occurs when a single nucleotide (A, T, C or G) in the genome sequence is altered (mutated). For example, a mutation might change the DNA sequence AAGGCTAA to ATGGCTAA.

The only difference between a point mutation and a SNP is that for a mutation to be classified as a SNP, all alternative forms must each occur with a frequency of 1% (0.01) or greater in a population.

Therefore, a point mutation is regarded as a di-allelic SNP when designing assays using Assay Design software.

Insertion/deletion polymorphisms (InDels)

An InDel is a mutation where one or several nucleotides have been inserted/deleted in a DNA sequence.

Use square brackets [] when entering InDels in Assay Design Software. It does not matter if the InDel is an insertion or deletion when entering it in the software (see Example 1).

If the InDel is located in a homopolymeric stretch, see Hints & Tips for more information.

Example 1:

A[C]G represents the two alleles **ACG** and **AG**

GC[TT]TT represents the alleles **GCTTTT** and **GCTT**

AT[AT][AT] can be used to represent a short tandem repeat with alleles **AT**, **ATAT**, and **ATATAT**. Select [AT][AT] as target region.

Short tandem repeats (STRs)

A short tandem repeat (STR) is a repeated DNA sequence in which the repeat elements are typically two to five base pairs long.

To design an assay for STR analysis, use the InDel notation (see **Example 2**).

Example 2:

Allele 1: ACGACGACG

Allele 2: ACGACGACGACG

Allele 3: ACGACGACGACGACGACG

Assay Design Software will not position primers inside the selected target region. In order for the sequencing primer to overlap the constant repeat region (when possible), two different notations and choices of target region should therefore be used for forward and reverse sequencing primers respectively.

Forward assay:

ACGACGACG[ACG][ACGACG]

The target region should be selected starting four nucleotides into the constant part of the repeat region. In the example above, choose CGACG[ACG][ACGACG] as target region and generate only forward sequencing primers (Tick the box "Generate forward primers" in the Current assay settings dialog). By choosing this target area, the forward sequencing primers that are generated will overlap between 0 and 4 nucleotides of the constant repeat region.

Reverse assay:

[ACGACG][ACG]ACGACGACG

The target region should be selected starting four nucleotides into the constant part of the repeat region. In the example above, choose [ACGACG][ACG]ACGAC as target region and generate only reverse sequencing primers (Tick the box "Generate reverse primers" in the Current assay settings dialog). By choosing this target area, the reverse sequencing primers that are generated will overlap between 0 and 4 nucleotides of the constant repeat region.

11.1.3 Entering polymorphisms in Assay Design Software

The following rules apply when entering polymorphic positions in the software:

- **Single nucleotide polymorphisms (SNPs)** should be entered using "/" (e.g. A/T) or by using the corresponding IUPAC code.

IUPAC code	Refers to the following nucleotides
A	Adenine
C	Cytosine
G	Guanine
T	Thymine
M	A/C
R	A/G
W	A/T
S	G/C
Y	T/C
K	G/T
V	A/C/G
H	A/C/T
D	A/G/T
B	C/G/T
N	G/A/T/C

- **Insertion/deletion polymorphisms (InDels)** should be entered using the square bracket notation "[]" (e.g. [AT]). **Short Tandem Repeats (STRs)** should be entered using the InDel notation. See **section 9.9** for further information.
- Polymorphisms involving a combination of SNPs and InDels should be entered using a combination of "/" and "[]" (see example 3 below).
- Nested polymorphisms are not allowed. Example: [ATT[C]G].

The table below shows some examples of how to enter different kinds of polymorphisms in the software.

Example 3:

SNPs	InDels	SNPs and InDels in combination
<p>Di-allelic SNP: G/C or S (IUPAC)</p> <p>Tri-allelic SNP: A/C/G or V (IUPAC)</p> <p>Tetra-allelic SNP: G/A/T/C or N (IUPAC)</p>	<p>A[C]G: represents either ACG or AG</p> <p>GC[TT]TT: represents either GCTTTT or GCTT</p>	<p>[T/A] represents a tri-allelic polymorphism where the possible alleles are a T, an A or neither (deletion).</p>

11.2 Sequence analysis (SQA)

11.2.1 Introduction

Assay Design Software can be used to design assays (PCR and sequencing primers) for sequence analysis (SQA) in the PSQ 96MA System.

This part of the methodology provides information on how to enter the sequences for sequence analysis in the Assay Design Software.

11.2.2 Entering sequences for SQA into Assay Design software

Sequence containing variable region with unknown variants

Enter, or import, the sequence in **Sequence Editor** with a number of N-nucleotides to denote the unknown sequence.

E.g. ACCAGTATTTAGGACCAGATTAGGNNNNNNNNNNNNNNNNNNNNNNNACCAGGATGACAGTAGACCC

The default target region will then be the stretch of unknown nucleotides with three known, non-variable nucleotides flanking the unknown region on either side.

Default target region:

ACCAGTATTTAGGACCAGATTAGGNNNNNNNNNNNNNNNNNNNNNNNACCAGGATGACAGTAGACCC

Sequence containing variable region with known variants

1. Enter the sequence of the expected, or most common, variant.

Expected sequence variants:

ACGTGGCTGG CATGGCTGCT CG (ACGTGGTG/ACGTGCTG/ACCCTCC)
CGTGCTGCAT GGCTGCGCTG CTG

Enter:

ACGTGGCTGG CATGGCTGCT **CG ACGTGGTG CGT**GCTGCAT GGCTGCGCTG CTG

2. Set the nucleotides highlighted in blue as target region (i.e. the variable region flanked by three known, non-variable nucleotides on each side).

Confirming a known sequence

Enter the expected sequence, and mark the region of interest as the target region.

12 Glossary

A

Adenine (A): A purine base that is a part of DNA and RNA molecules. Adenine forms base pairs with thymine (a pyrimidine base) and uracil (in the case of a RNA molecule).

Allele: An allele is any one of a number of alternative forms of the same gene occupying a given locus (position) on a chromosome.

Allele frequency: The estimation of the proportion of each allele at one gene locus (for example, the proportion of each allelic variant in a SNP) in a population.

Allele quantification: Allele quantification is used to estimate the allele frequencies for sample populations.

Amplicon: The amplicon is the DNA sequence amplified in a PCR reaction and is defined by the PCR primers on either side of the target region (forward and reverse primer).

Annealing: The base pairing of a primer to a complementary single strand of DNA (or RNA) to form a double helix. The bases are held together by hydrogen bonds formed during annealing.

Annealing over a SNP: The user can allow the PCR- and/or sequencing primers to anneal over a SNP, if necessary. This is useful if a polymorphism needs to be genotyped where the surrounding sequence contains SNPs.

Assay: An assay is defined as the information required to perform PCR amplification and Pyrosequencing analysis on a sample. The unique feature of an assay is the target region, which either contains the polymorphism(s) of interest (in genotyping or AQ) or the (unknown) SQA sequence. An assay generated by Pyrosequencing Assay Design Software contains information on suitable PCR primers, sequencing primers and corresponding sequences to analyze.

B

Base: A base is a part of nucleotides, which are the building blocks of DNA. Adenine, thymine, cytosine, and guanine are bases. Adenine forms base pairs with thymine and cytosine forms base pairs with guanine. They are called bases because they are alkaline (basic) in the acidic DNA structure.

Base pair: Represents two complementary bases bound together by hydrogen bonds. In DNA, adenine (A) is hydrogen bonded to thymine (T) and guanine (G) is hydrogen bonded to cytosine (C). Two strands of DNA bound together by base pairs form a double helix. The number of base pairs is often used as a measure of the length of a DNA segment.

Batch assay design: Simultaneous assay design on several DNA sequences.

Biotin: A molecule that can bind very strongly to streptavidin. Can be used for biotinylation of primers to bind a DNA strand to a streptavidin-coated solid phase.

BLAST homology search: BLAST (Basic Local Alignment Search Tool) is an algorithm that can be used to search sequence databases for homologous sequences. It may be used to characterize unknown sequences, or to find homologous or related sequences.

C

Chromosome : A physically distinct unit of the genome containing many genes. The chromosomes are replicated (duplicated) during cell division. Prokaryotic genomes often carry the entire

genome on one circular chromosome whereas eukaryotic genomes often have a number of chromosomes.

Cyclic dispensation order: A repetitive dispensation order for nucleotide dispensation in the Pyrosequencing reaction. Normally used in Pyrosequencing technology for sequencing unknown DNA-sequences. For example, "CTGA" or "TCGA" can be used and repeated the desired number of times.

Cytosine (C): A pyrimidine base that is a part of DNA and RNA molecules. Cytosine forms base pairs with guanine (a purine base).

D

Deletion: Loss of a base or a segment of DNA from a chromosome. Small deletions within a gene can alter the reading frame, and thus the amino acid sequence of the encoded protein.

Deoxyribonucleic acid (DNA): The carrier of genetic information in most organisms.

Di-allelic SNP: A SNP where two possible bases can occupy the polymorphic position.

Diploid: A diploid genome contains two copies of each chromosome.

Directed dispensation order: Non-cyclic nucleotide dispensation order that follows the known sequence. To be used in Pyrosequencing technology when you know the sequence to be analyzed. Example: the sequence TCCAGAA should be dispensed TCAGA.

Direction: The direction of an assay, which can be either forward or reverse. See **Forward assay** and **Reverse assay** respectively.

Dispensation order: Defines the order in which nucleotides should be dispensed in a Pyrosequencing run. Use a cyclic dispensation order if the sequence is unknown. Use a directed dispensation order if analyzing a SNP, or if the sequence is known.

DNA pool: A mix of several DNA samples.

Duplex: A double-stranded nucleic acid, formed e.g. by two primers that anneal to each other.

E

Enzyme: A protein (or RNA) working as a catalyst to enhance the speed of a biochemical reaction without altering it.

EMBL Nucleotide Sequence Database: The EMBL Nucleotide Sequence Database is Europe's most comprehensive nucleotide sequence database

F

FASTA format: The FASTA file format is a common format for DNA sequence files. A sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The description line is distinguished from the sequence data by a greater-than (>) symbol in the first column. It is recommended that all lines of text are shorter than 80 characters in length.

Forward assay: A forward assay is an assay where the sequencing primer is annealed to the complementary strand of the input DNA sequence. Sequencing will be performed, reading the polymorphism/unknown region in forward direction, i.e. generating the input sequence. For a forward Pyrosequencing assay, the reverse PCR primer should be biotin-labeled.

Forward PCR region: The area in the DNA sequence in which forward PCR primers will be placed when performing primer set generation.

G

GC content: The percentage of G and C bases in a DNA molecule.

GenBank: GenBank is an annotated collection of all publicly available DNA sequences and the genetic sequence database of the US National Institute of Health (NIH).

Gene: A specific DNA sequence from which a protein or RNA can be generated. The human genome is estimated to contain 30,000 genes.

Genome: The total genetic information of an organism.

Genomics: The analysis and investigation of the genome.

Genotype: The observed alleles at a genetic locus for an individual.

Gibb's free energy (ΔG): Gibbs free energy is a measurement of nucleic acid duplex stability. A DNA duplex is more stable when its ΔG value is more negative. The definition of free energy is: $\Delta G = \Delta H - T\Delta S$, where H is the enthalpy, S is the entropy, and T is the temperature.

Guanine (G): A purine base that is a part of DNA and RNA molecules. Guanine forms base pairs with cytosine (a pyrimidine base).

H

Hairpin: See **Hairpin loop**.

Hairpin loop: A self-annealing nucleic acid strand, forming stable hydrogen bonds ($\Delta G < \text{zero}$) with itself. The hydrogen-bonded region is referred to as a stem and the single-stranded region is referred to as a loop.

Heterozygote: An individual with different alleles at a given locus (position) on the two corresponding chromosomes.

Homopolymer: A stretch of identical bases in DNA. In Pyrosequencing technology, stretches of two or more identical bases are regarded as homopolymers.

Homozygote: An individual with the same alleles at a given locus (position) on the two corresponding chromosomes.

I

Insertion: Addition of a base or a DNA segment into a chromosome. Small insertions within a gene can alter the reading frame, and thus the amino acid sequence of the encoded protein.

International Union of Pure and Applied Chemistry (IUPAC): An organization providing recommendations on organic and biochemical nomenclature, symbols, terminology, etc.

L

Locus: The position on the chromosome at which a gene or a genetic marker is situated. The locus may be occupied by any one of the alleles for the gene or the genetic marker.

M

Melting temperature: Melting temperature (T_m) of a primer is defined as the temperature at which 50% of the primer is annealed to the template, and 50% is free in the solution. The melting temperature depends on the primer sequence (GC-rich primers have higher melting temperatures), primer length, concentration, and chemical properties of the buffer

solution. There are different ways to calculate the melting temperature. The method primarily used in the Assay Design Software is the Nearest neighbor method.

Mispriming: Mispriming occurs when there are alternate annealing sites for the primer in the template sequence, i.e. when the 3'-end of a primer has significant homology with more than one site on the template sequence.

Multiplex assay: Reaction with one or several different DNA templates and several sequencing primers in the same reaction, one primer per polymorphism.

Mutation: A change in the DNA sequence. Mutations within a gene can alter the amino acid sequence of the encoded protein. They can also alter the reading frame, and thus the amino acid sequence of the encoded protein. Mutations occurring more frequently than 1% in a population are SNPs.

N

Nearest Neighbor algorithm: A common method for calculating the melting temperature (T_m) of primers.

Nucleoside: A compound consisting of a purine or pyrimidine base covalently linked to a pentose.

Nucleotide: A nucleoside phosphorylated on one or more of the hydroxyl groups of the sugar. A nucleotide is the monomer unit of nucleic acids.

P

PCR: See **Polymerase chain reaction**.

PCR primer: A primer used for amplifying a part of a DNA sequence by Polymerase chain reaction (PCR).

PCR primer pair: In order to conduct PCR, two primers are necessary, one forward and one reverse. Hence PCR primers always come in pairs.

PCR product: See **Amplicon**.

PCR region: The forward PCR region is the region in the DNA sequence where the forward PCR primer will be positioned, and the reverse PCR region is the region in the DNA sequence where the reverse PCR primer will be positioned.

Point mutation: A variation in a single nucleotide position. The most common type of genetic variation. Point mutations that occur in more than 1% of the members of a population are called SNPs.

Polymerase chain reaction (PCR): A technique for amplifying a specific segment of DNA more than 1 million times. Multiple cycles of denaturation, annealing with primer, and extension using a thermostable DNA polymerase produce an analyzable amount of DNA.

Polymerases: Enzymes that catalyze the synthesis of nucleic acids, assembling DNA or RNA from (deoxy)ribonucleotides. Most polymerases need one strand of pre-existing nucleic acid as template and a double-stranded end to begin from.

Polymorphism: Genetic variations, broadly encompassing any of the many types of variations in DNA sequence that are found within a given population. Specific subtypes of polymorphisms include mutations, point mutations, SNPs and insertions/deletions.

Primer: A short DNA oligonucleotide that anneals to a template DNA strand. The primer provides a free 3'-OH end from which the DNA polymerase can start synthesizing a complementary DNA strand. Primers can be labeled with various molecules, e.g. biotin.

Primer annealing site: The primer annealing site is the location on the template DNA strand where the primer can anneal. In general, the primer has been designed to be 100% complementary to the template DNA strand, which allows specific primer annealing.

Primer set: A primer set for Pyrosequencing consists of a PCR primer pair and a sequencing primer.

Pyrogram: The resulting graph from a sequencing reaction performed using Pyrosequencing technology. Each incorporated nucleotide is shown as a peak in the pyrogram.

R

Reference peak: Reference peaks are used as internal controls in sequencing reactions in Pyrosequencing technology.

Repeat: See **Short tandem repeat**.

Reverse assay: A reverse assay is an assay where the sequencing primer is annealed to the input DNA sequence strand. Sequencing will be performed reading the polymorphism/unknown region in the reverse direction, i.e. generating the complementary strand to the input DNA sequence. The forward PCR primer should be biotin-labeled for a reverse Pyrosequencing assay.

Reverse PCR region: The area in the DNA sequence in which reverse PCR primers will be placed when generating primer sets.

S

Secondary structures: Structures formed by single-stranded DNA, such as hairpins and loops. Often seen in GC-rich or repetitive stretches of the DNA. Secondary structures may be difficult for polymerases to read through.

Self-priming: Self-priming occurs when a part of the DNA sequence at the 3'-end is complementary to another part of the DNA sequence and the strand binds to itself and acts as a primer. In the Pyrosequencing reaction, self-priming may result in an overlaid sequence, making the pyrogram difficult to interpret.

Sepharose beads: Streptavidin-coated sepharose beads that can be used for preparation of biotinylated PCR products.

Sequence format: The way in which a DNA sequence is recorded in a computer file. Different programs for database searches use different formats. Examples of formats are FASTA, GenBank and EMBL formats

Sequence to analyze: A short part of a DNA sequence, starting directly after the sequencing primer, which contains one or several polymorphisms to be analyzed using PSQ 96MA or PSQ HS96A systems. See also "The Sequence to Analyze" in the methodology part of the manual.

Sequencing primer: A primer used for sequencing part of a DNA sequence.

Short tandem repeat (STR): An STR is a polymorphism where two or more bases are repeated after each other with different (limited) multiplicities in different alleles.

Simplex assay: Sequencing reaction with one sequencing primer from which one, or several, polymorphisms are analyzed in a reaction. Compare with **Multiplex assays**.

Single nucleotide polymorphism (SNP): SNPs involve the change of one DNA base to another. SNPs and point mutations are structurally identical, differing only in their frequency. Variations that occur in 1% or less of a population are considered point mutations, while those occurring in more than 1% are SNPs. SNPs can occur in coding regions of the genome (cSNPs), in regulatory regions (rSNPs), or, most commonly, in "junk DNA" regions, in which case they are referred to as anonymous SNPs. SNPs can be di-, tri-, or tetra-allelic polymorphisms. However, in humans, tri-allelic and tetra-allelic SNPs are rare.

STR: See **Short tandem repeat**.

Streptavidin: A molecule that binds very strongly to biotin.

T

Target region: The target region defines the region to be analyzed by Pyrosequencing. In genotyping and AQ, the target region will contain the polymorphism(s) that will be genotyped. In SQA, the target region will contain the known or unknown sequence to be sequenced. Primers generated by Pyrosequencing Assay Design Software will always be placed outside the defined target region.

Template: The DNA template is commonly defined as the (single-stranded) DNA that is used by the DNA polymerase to synthesize a complementary "template copy". In Pyrosequencing, the template is specifically the single stranded DNA that is attached to the Sepharose beads (by means of a biotin bridge) in the well, and which is used to perform sequencing-by-synthesis.

Tetra-allelic SNP: A SNP where any one of four bases can occupy the polymorphic position.

Thymine (T): A pyrimidine base that is a part of DNA molecules. Thymine forms base pairs with adenine (a purine base).

Tri-allelic SNP: A SNP where any one of three bases can occupy the polymorphic position.

13 Index

Amplicon length	47	PCR primer end stability	46
Analysis steps	42	PCR primer length	46
Annealing over a SNP	80	PCR primer pair analysis	47
Assay	80	PCR primer regions	28
Assay controls	60	Performing an assay design	16
Assay design settings for one assay	12	Polymorphism	49
Assay overview area	11, 33	Polymorphisms area	11, 18, 26, 27
Assay settings	34	Primer set analysis	49
Assay setup	17	Print a report	38
Assay type	17	Printer	7
Assay window	11, 17, 33, 36	Processor	7
Batch assay design	39, 80	Quality	31
Biotinylated PCR Primer Hairpins	49	RAM	7, 65
Computer	7	Report	33, 37
Copy all primer sets	33	Report formats	37
Copy and paste a DNA sequence	18	Result presentation	31, 33, 37
Copy entered sequence	18	Reuse primers	29
Copy primer set	33, 61	Reverse assay	81, 84
Default assay design settings	12	Reverse PCR region	84
Discard quality	32	Run automatic design	28
Dispensation order	41, 72, 81	Save a report	38
Edit primers	35	Save an assay	36
EMBL Nucleotide Sequence Database	20, 81	Score	31
Enter the DNA sequence	17	Scoring	31
FASTA format	18, 19, 71, 72, 81	Search	27
Final primer set	34	Seq primer complementarity	48
Find sequence	27, 29, 30	Seq primer duplexes	48
Forward assay	81	Seq primer Hairpin loops	48
Forward PCR region	81	Seq primer length	48
GC content	46, 48	Seq primer melting temperature	48
GC content difference	47	Sequence Editor tab	17
GenBank	23, 82	Sequence tab	33
Hard drive	7	Sequence to analyze	49, 72
High quality	32	Sequencing primer analysis	48
Homopolymers	48	Settings	12
Import a DNA sequence	18, 19	Buttons	14
Import into PSQ software	41	PCR Primer settings	12
InDels	63	Primer set settings	13
Installation	7	Sequencing primer settings	13
License key	10	Short tandem repeats	64, 77
Low quality	32	Simplex assay	73
Medium quality	32	Software setup	7
Melting temperature	46	Sort primers	32
Melting temperature difference	47	System requirements	7
Menu bar	11	Target distance	48
Monitor	7	Target region	26, 27
Multiplex assay	61, 73, 74, 83	Template loops	49
Name polymorphisms	26	The Assay Design Software start screen	11
Operating system	7	Troubleshooting	65
PCR Complementarity	46	Type in a DNA sequence	18
PCR duplexes	46, 47	Universal biotinylated PCR primers	61
PCR hairpin loops	46	Warning messages	44
PCR Mispriming	46	Warnings and penalties	42
PCR primer analysis	46	Work area	11

