NSA-IPI: User Manual

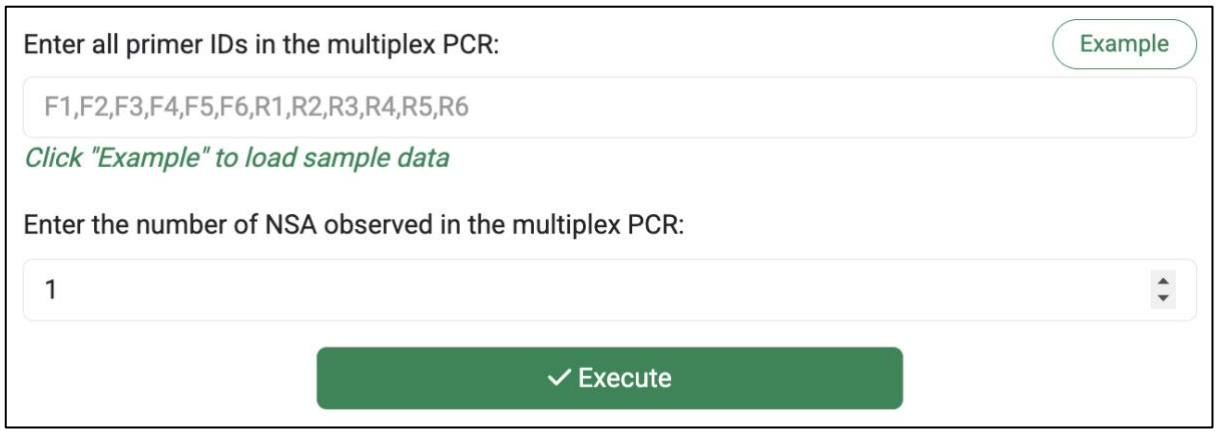
This manual provides a detailed guide for using NSA-IPI (Non-Specific Amplification Implicated Primer Identification), an online tool designed to help researchers identify the specific primer pairs responsible for non-specific amplifications (NSAs) in multiplex PCR.

Module 1: NSA-Implicated Primer Identification

This module uses a combinatorial algorithm to redistribute inputted primers into four diagnostic pools (A, B, C and D), allowing the user to identify problematic primer pairs with a minimal number of investigatory PCR reactions.

Step 1: Initial Input and Primer Pool Generation

- **Input Primer IDs:** Enter all primer IDs from the original multiplex PCR where the NSA bands were observed.
- **Specify NSAs:** Enter the number of observed NSA bands on the gel electrophoresis visualisation of the original multiplex PCR.
- **Generate Diagnostic Primer Pools:** Click **"Execute"** to have the NSA-IPI tool generate four primer pools (A-D).
    - *Note: Click **"Example"** to load sample data, then click **"Execute"** to view a demonstration of the output.*



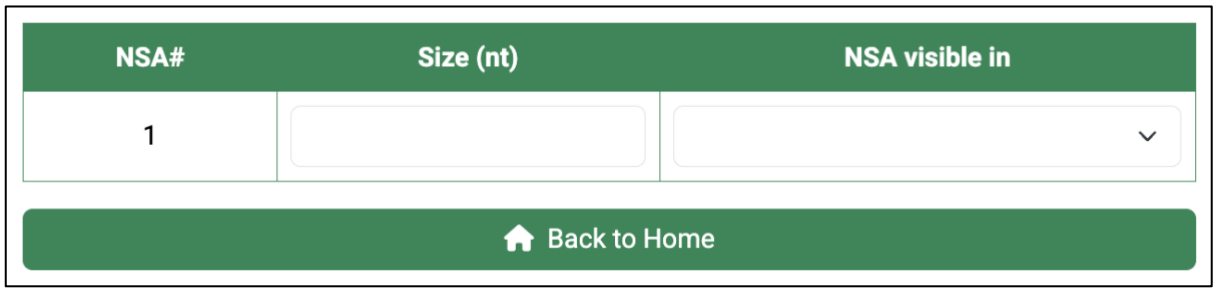
Section of the NSA-IPI input page for the "NSA-Implicated Primer Identification" module, showing the initial entry fields (i.e., Primer IDs and Number of observed NSAs), as well as the **"Example"** and **"Execute"** buttons.

Step 2: Investigatory PCR Assay

- **Conduct PCR:** Perform multiplex PCR using the generated four primer pools (A-D).
- **Perform Gel Electrophoresis:** Visualise the PCR products of these diagnostic pools (A-D) via gel electrophoresis to see which pools still produce the NSA bands.

Step 3: Candidate NSA-Causing Primer Pair List

- **Input NSA Size:** Enter the size of the target NSA in the **"Size (nt)"** field on the output page where the suggested A-D primer pools are displayed. This facilitates differentiation when multiple NSA bands are present.
- **Select NSA Lanes:** Use the **"NSA visible in"** dropdown menu to select which lanes show the NSAs, based on the gel electrophoresis visualisation of the diagnostic A-D pools.
- **Analyse Candidate Primer Pair List:** If the NSA-IPI generated NSA-causing primer pair list is small, proceed with confirmatory singleplex PCR. However, if the list is still large, run another round of NSA-IPI analysis on the resulting item list.



Section of the NSA-IPI output page for the "NSA-Implicated Primer Identification" module, showing the **"Size (nt)"** entry field and the **"NSA visible in"** dropdown menu, as well as the **"Back to Home"** button.

Step 4: Actions for NSA Resolution

Once the NSA-causing primer pairs are identified, there are two options to eliminate the NSAs:

- **Separation:** Move the problematic primers into different PCR reaction tubes, ensuring they never interact in the same multiplex reaction.
- **Redesign Mis-annealing Primers:** Use the **"Mis-Priming Primer(s) Prediction"** module (detailed below) to identify the specific mis-annealing primers within the problematic pairs for redesign.

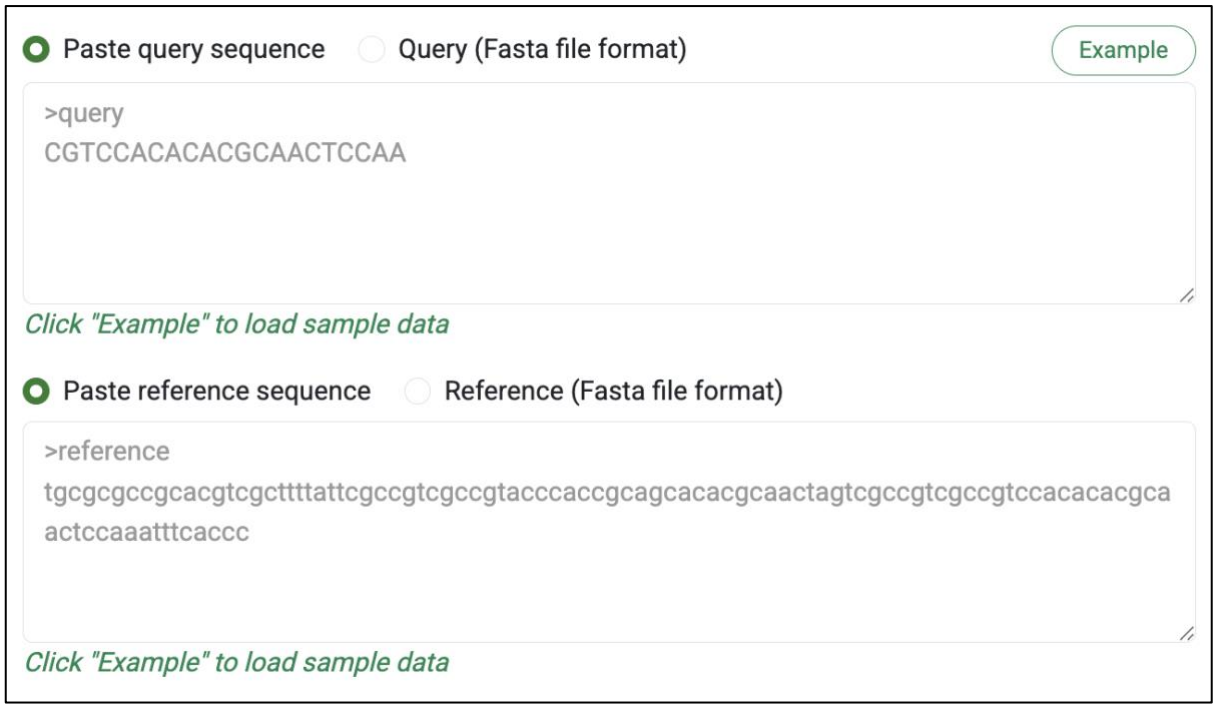
Module 2: Mis-Priming Primer(s) Prediction

This module helps the user to determine which specific primers within the NSA-causing pairs identified by NSA-IPI are binding incorrectly (mis-annealing).

Step 1: Preparing the Sequence Data

Analyse each primer of the identified NSA-causing pair independently to compare their alignment strengths and, consequently, determine the mis-annealing primer(s).

- **Query Input:** The reference of the “assumed” mis-annealing primer (Fasta format)
- **Reference Input:** A short sequence from the reference genome covering the “assumed” mis-annealing site (Fasta format).
    - **Defining the Short Reference Sequence:** This is determined based on the length of the NSA of interest and the specific binding site of the **other** primer in the NSA-causing pair (NOT assumed to be mis-annealing).
    - **Scenario A (Testing the Reverse Primer):** If the Forward primer is assumed to be binding specifically, retrieve the reference sequence from a location downstream (3') of that Forward primer at an offset equal to the NSA length.
    - **Scenario B (Testing the Forward Primer):** If the Reverse primer is assumed to be binding specifically, retrieve the reference sequence from a location upstream (5') of that Reverse primer at an offset equal to the NSA length.



Section of the NSA-IPI input page for the “Mis-Priming Primer(s) Prediction” module, showing the **“Query”** and **“Reference”** entry fields, as well as the **“Example”** button.

Step 2: Configuring and Running MegaBLAST

The default MegaBLAST settings within NSA-IPI enable detection of mis-annealing events during sequence alignment analysis. However, these parameters can be optimised for more sensitive detection of weak mis-annealing events.

- **Initial MegaBLAST Analysis:** Perform sequence alignment analysis for the NSA-causing pair using the default MegaBLAST settings.
    - **MegaBLAST Parameter Adjustment:** If "No hits found" result occurs for **both** primers following this initial analysis, increase the MegaBLAST sensitivity by:
        - Decreasing the Word Size (e.g., from 7 down to 5).
        - Reducing Gap Penalties (Gap Open/Extend).
        - Increasing the E-Value threshold.
- **Run MegaBLAST:** Click **"Run BLAST"** button to execute MegaBLAST-powered sequence alignment analysis.
    - *Note: Click **"Example"** to load sample data, then click **"Run BLAST"** to view a demonstration of the output.*

Step 3: Interpreting the Output

- **MegaBLAST Output Evaluation:** Compare the results for both primers in the pair using these metrics:
    - **Alignment Directionality:** Check if the alignment orientation (e.g., 5' to 3') matches the expected role of the primer in producing that specific NSA. The tool displays this orientation as a (Reference strand)/(Query strand) relationship, where a "Plus/Plus" (+/+) result indicates the primer under investigation is acting as a Forward primer, while a "Plus/Minus" (+/-) result indicates it is acting as a Reverse primer.
    - **Hit Score (Bit Score):** A higher score indicates a stronger, more likely mis-annealing event.
    - **E-Value:** Lower values indicate the alignment is less likely to have occurred by chance.
- **Verdict:** If one primer shows a significantly higher hit score and lower E-value than the other, and its alignment orientation is consistent with the expected role in producing the observed NSA, it is likely the primary source of mis-annealing. However, if both primers return very low hit scores or "No hits found" results despite relaxed MegaBLAST settings, both may be mis-annealing.