Introduction

The evolution of high-throughput molecular tools such as DNA microarrays or gene capture methods coupled with next-generation sequencing opened the door to unprecedented opportunities in microbial ecology to appraise the taxonomic and functional diversity of microorganisms within complex environments. The success of these high-throughput strategies, however, strongly relies on the quality of designed probes. Besides molecular applications, these probes can be used in silico to provide a fast and reliable annotation of massive sequencing data. Consequently, probe design is of critical importance and therefore multiple parameters should be considered in order to ensure the selection of sensitive, specific and explorative probes.

The KASpOD (k-mer based Algorithm for high-Specific Oligonucleotide Design) software has been developed to design such probes based on large datasets. This multipurpose tool was implemented to design probes from the exponentially growing sequence datasets in microbial ecology by using highly parallel computing architectures and an innovative k-mers based strategy that allowed overcoming major limitations in this field.

KASpOD Algorithm

- **Coverage assessment**
  - Target group
  - Non-target group
- **Probe design**
  - k-mers
  - Evaluate specificity and affinity
- **Evaluation**
  - Group-specific k-mers
  - Clustering
  - Generate oligonucleotide probe
  - Elute on magnetic beads

Web Service

http://giim.u-clermont1.fr/kaospod

Applications

- **(A) Gene capture by hybridization**
  - Innovative strategy for microbial ecology
- **(B) Environmental microarrays**
  - Efficient exploration of microbial diversity
  - Probe design based on the Greengenes (4) 16S rRNA sequence database
  - 1 295 prokaryotic genera
  - 22 613 degenerate signatures
- **(C) PhyLOPDb**
  - 16S rRNA oligonucleotide probe database
  - 74 003 25-mer probes targeting 2 178 prokaryotic genera
  - Applications: PCR, FISH, gene capture, in silico rapid sequence identification
- **(D) AFFILGOOD strategy**
  - Efficient algorithm to perform taxonomic and functional assignment of metagenomic reads
  - High-throughput
  - Competitive results

References


†: these authors contributed equally to this work.

Contact

Nicolas Parisot: nicolas.parisot@insa-lyon.fr
Eric Peyretaillade: eric.peyretaillade@udarmail.fr