Základy bioinformatiky
Tutorial 3: Design PCR primerů

Pojem primer se v molekulární biologii používá pro krátký oligonukleotid, který slouží jako začátek synthesy DNA prostřednictvím DNA-polymerasy. Pojem se používá jak pro organismem vytvořený počátek *in vivo* replikace DNA, tak i pro synthetický oligonukleotid, který slouží jako základ *in vitro* polymerasové řetězové reakce (PCR). Krátké oligonukleotidy požadované sekvence je dnes možné získat chemickou synthesou, nejčastěji na zakázku. U zakázkových výrobců oligonukleotidů je možné na webových stránkách zadat požadované sekvenci, množství, čistotu, případně modifikace oligonukleotidu a pak jej stisknutím tlačítka objednat.

PCR je možné použít buď pro detekci nebo pro isolaci určitého úseku DNA, případně pro jiné speciální účely. V tomto tutoriálu si ukážeme obě varianty, jak detekci, tak i isolaci úseku DNA. Nejprve si navrhneme primery pro detekci genu první podjednotky mamutí cytochrom-c-oxidasy. Nejčastější použití PCR vyžaduje dvojici primerů, které předurčují produkt reakce. Pokud máme k dispozici nukleotidovou sekvenci úseku, který chceme detekovat, psanou od 5' k 3' konci, pak první primer můžeme vybrat jako úsek sekvence někde na začátku sekvence. Druhý primer můžeme vybrat tak, že vezmeme nějaký úsek na konci sekvence a přeložit jej do komplementární sekvence. Požadavky na primery se mohou lišit podle toho, jestli se jedná o klasickou PCR nebo kvantitativní PCR s reverzní transkripcí, na použité DNA-polymerase a na dalších aspektech. Zde uvedené požadavky by měly fungovat pro většinu použití PCR, ale pro konkrétní aplikaci PCR si zjistěte konkrétní specifikace. Podle obecných podmínek by měl primer mít **18 až 24 b**, teplotu tání T_m **52-58 °C** a **obsah G-C párů 40-60 %** . Teplota tání se počítá podle jednoduchého vzorečku. Neměl by být velký rozdíl teplot tání ve dvojici primerů. Dále jsou vítány páry GC na 3' konci (mezi posledními 5 b). Sekvence by neměla obsahovat mnoho repetic (maximálně 4 pro jednu basi). Výtěžek PCR může být negativně ovlivněn přítomností sekundárních struktur v primeru a schopností primeru tvořit dimer sám se sebou nebo s druhým primerem. Primery je potřeba vytvářet tak, aby bylo dosaženo vhodné délky produktu, která činí cca 100 bp pro kvantitativní PCR a 100-500 bp pro konvenční PCR s použitím obyčejné DNA-polymerasy. Pro amplifikaci delších úseků je nutné použít speciální DNA-polymerasu s opravným mechanismem. V neposlední řadě je potřeba vybrat primery tak, aby byly co nejvíce specifické a v daném vzorku amplifikovaly pouze požadovanou sekvenci.

Návrh primerů si ukážeme na nástroji *Primer BLAST* z nabídky nástrojů *National Center for Biotechnology Information* (NCBI). Nejprve si na stránkách http://www.ncbi.nlm.nih.gov/ najdeme sekvenci mamutí cytochrom-c-oxidasy zadáním jejího UniProt identifikátoru **`COX1_MAMPR`**. Když kliknete na identifikátor, objeví se vám záznam proteinové sekvence. My ale chceme nukleotidovou sekvenci. Tu dostaneme kliknutím na záznam **`DQ188829.2`**. Tím se vám objeví nukleotidová sekvence kompletního mamutího mitochondriálního genomu s identifikátorem **`DQ188829`**. Postupným rolováním dolu najdeme úsek odpovídající cytochrom-c-oxidase. Odpovídající nukleotidová sekvence se vám zvýrazní kliknutím na blízký odkaz *CDS* (coding DNA sequence). Sekvenci zkopírujte a vložte do nástroje Primer BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Dále v části *Primer Pair Specificity Checking Parameters* nastavte databasi *nr* (non-redundant sequences) a smažte *Homo sapiens* v položce *Organism*. Díky tomu bude program zjišťovat, jestli do jaké míry budou primery reagovat s dalšímu sekvencemi v databasi *nr*.

Po zmáčknutí *Submit* se vám objeví několik navržených dvojic primerů. Pro každou dvojici se vám objeví délka produktu PCR reakce. Pro každý primer získáte délku, jeho počátek a konec v sekvenci, teplotu tání, podíl GC párů a informace o tvorbě sekundárních struktur a dimerů. Dále vám nástroj nalezne sekvence v databasi *nr*, na které by fungoval. Tím můžeme posoudit specificitu primeru. Pokud byste vyvíjeli primer, který by měl detekovat například přítomnost viru ve vzorku lidské krve, pak by bylo vhodné místo database *nr* zvolit lidský genom a tak otestovat, jestli by primer nefungoval i na nějakou lidskou sekvenci.

Jiným způsobem jak využít program *Primer BLAST* je v našem případě zadání sekvence celé mitochondriální DNA (nebo jejího kódu **`DQ188829`**) do políčka sekvence. Aby program věděl, že má hledat primer jen pro cytochrom-c-oxidasy, zadejte do položek *Range* hodnoty 5331 pro začátek forward primeru a 6878 pro konec sekvence reversního primeru. Výhodou je fakt, že program vezme v úvahu zbytek sekvence a navrhne primery tak, aby nenasedaly na jiné části sekvence.

Další možností je nezadávat žádnou sekvenci, zadat jen sekvence primerů a kterou databasi chceme prohledávat. Program nám nalezne podobné sekvence v databasi *nr*, případně jiné kterou vybereme. Toto použití si můžete vyzkoušet tak, že jako primery použijete sekvence nalezené v předchozích krocích. Stejně využijeme nástroj i pro nalezení primeru pro amplifikaci konkrétního úseku DNA. Například pokud bychom chtěli rekombinantně exprimovat mamutí cytochrom-c-oxidasu v bakterii *E. coli*, pak by bylo potřeba amplifikovat celý gen. Podívejte se na sekvenci mitochondriální DNA mamuta s vyznačenou sekvencí kódující cytochrom-c-oxidasy. Jako forward primer vyberte například prvních 20 basí genu, tedy **`atgtttgctaaccgctgact`**. Reverse primer získáte tak, že vyberete sekvenci posledních 20 basí genu (**`acgtaaaatctaactcgaga`**) a vytvoříte k ním komplementární sekvenci (nestačí jen změnit **`a`** na **`t`** atd, ale ještě musíte napsat sekvenci pozpátku). Když obě sekvence zadáte do nástroje *Primer BLAST*, pak vám program poskytne kromě nalezených vazebných sekvencí navíc informace o teplotě tání, obsahu GC a podobně. Sekvence je možné upravit, například pokud je nízká teplota tání, tak primer prodloužit. Program ještě obsahuje volby týkající exonů a intronů. Pokud chcete primer analysovat aniž byste prohledávali sekvence programem *BLAST*, pak je možné buď tuto volbu vypnout, nebo použít jiný program, například http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/.