

# UŽITNÝ VZOR

(11) Číslo dokumentu:

# 34 717

(13) Druh dokumentu: **U1**

(51) Int. Cl.:

*C12Q 1/68* (2018.01)

*C12Q 1/6876* (2018.01)

*C12Q 1/6895* (2018.01)

(19)  
ČESKÁ  
REPUBLIKA



ÚŘAD  
PRŮMYSLOVÉHO  
VLASTNICTVÍ

(21) Číslo přihlášky: **2020-38246**

(22) Přihlášeno: **18.11.2020**

(47) Zapsáno: **22.12.2020**

(73) Majitel:  
ESSENCE LINE, s.r.o., Praha 5, Motol, CZ  
Zemědělský výzkum, spol. s r.o., Troubsko, CZ  
Masarykova univerzita, Brno, Brno-město, CZ  
Ing. Hana Jakešová, CSc., Hladké Životice, CZ

(72) Původce:  
Ing. Hana Jakešová, CSc., Hladké Životice, CZ  
Mgr. Ján Boroň, Prešov, SK  
Ing. Petr Novotný, Praha 5, Stodůlky, CZ  
prof. RNDr. Jana Řepková, CSc., Brno, Černá Pole, CZ  
RNDr. Jan Nedělník, Ph.D., Brno, Trnitá, CZ  
Ing. Oldřich Trněný, Rousínov, Slavíkovice, CZ  
Iker Silva Gonzales, Sopelana, Bizkaia, ES  
MSc. Mindaa Pauline, Saint-Faust, FR  
Mgr. Zuzana Čočková, Březová, CZ  
Mgr. David Vlk, Brno, Královo Pole, CZ  
Mgr. Petra Jirásková, Karlovy Vary, Drahovice, CZ

(74) Zástupce:  
Ing. Libor Markes, Grohova 145/54, 602 00 Brno,  
Veveří

(54) Název užitého vzoru:  
**Testovací sada k přípravě populačního čipu  
pro efektivní genotypování a selekci jetele  
lučního**

## Testovací sada k přípravě populačního čipu pro efektivní genotypování a selekci jetele lučního

### 5 Oblast techniky

Technické řešení se týká populačního čipu pro genotypování vybrané sady markerů a selekci jetele lučního (*Trifolium pratense* L.) prostřednictvím detekce alelových variant genů, využitelné zejména pro testování většího vzorku populace (řádově stovky až desetitisíce vzorků).

10

### Dosavadní stav techniky

15 Odvěkou snahou šlechtitelů je účinná selekce druhů rostlin a zvířat na základě fenotypové a nověji genotypové variace uvnitř druhu i populace. Tradiční šlechtění se dříve provádělo na základě fenotypu při absenci znalostí molekulárních a genetických mechanismů. Šlechtitelé prováděli selekci výběrem dle vybraných znaků jedinců pro příští potomstvo. Metodika šlechtitelů se prakticky nezměnila ani po objevech Mendelových zákonů genetiky (J. G. Mendel, 1866), ani rozšířením poznatků o DNA jeho následovníky ve 20. století (Morgan, Watson a Crick), takže 20 šlechtitelé zůstali odkázáni na selekci dle fenotypů.

O selekci jedinců na základě genotypování, tedy procesu určování genotypů pomocí laboratorních testů, je možno hovořit až s objevy moderních metod molekulární genetiky, zejména objevu 25 polymerázové řetězové reakce (PCR – polymerase chain reaction) v roce 1983 a Sangerova sekvenování (1977), které umožnily strukturní analýzu, detekci jednotlivých genů, genetických polymorfismů (existenci dvou nebo více variant genů, tzv. alel, v jednom lokusu) a chromozómových aberací.

Metody genotypizace zahrnují:

30

- identifikaci polymorfismu délky restrikčních fragmentů (RFLP – restriction fragment length polymorphism) genomové DNA,

35 - náhodnou amplifikaci polymorfni DNA (RAPD – Random Amplification of Polymorphic DNA) genomové DNA,

- detekci polymorfismu délky amplifikovaného fragmentu (AFLPD),

40 - polymerázovou řetězovou reakci (PCR),

- sekvenování DNA,

- sondy pro alelově specifický oligonukleotid (ASO),

45 - hybridizace na DNA populačního čipu či magnetických kuličkách (magnetic beads).

Do roku 1995, kdy byl popsán robotický tisk fragmentů DNA na skleněný substrát (Schena et al., 1995), bylo genotypování na jednotlivé i populační úrovni prováděno pouze na úrovni sériových 50 experimentů. Paralelní experimenty byly umožněny až s dostupností vysokokapacitních systémů (HTS), tj. nové generace sekvenování (NGS, Next-generation sequencing, přibližně od roku 2000) a technologie populačních čipů (MA Schena, US Patent 6,913,879, 2005). Právě populační čipy v současné době nacházejí uplatnění zejména ve šlechtění hospodářských zvířat a rostlin za účelem vyšší produkce potravin a zemědělských výnosů, jelikož představují robustní, ekonomický a efektivní nástroj pro genotypování živých organismů na populační úrovni. V synergii s využitím

nové generace sekvenování představují vhodnou platformu pro aplikaci dat získaných pokročilými sekvenačními metodami.

5 Pro genetiku a šlechtění různých zemědělských plodin a zvířat byly v minulosti vyvinuty a úspěšně použity různé typy molekulárních markerů, resp. fragmentů DNA. Jednotlivé genetické markery byly vyhodnoceny s různou úrovní reprodukovatelnosti, polymorfismů, genomické abundance; dále náročnosti, nákladnosti metodik potřebných k jejich detekci ad. Genetický marker je gen (nebo určitá sekvence DNA se známou lokací v chromozomu), který reguluje konkrétní gen nebo fenotypový znak. Mezi tyto markery patří polymorfismus délky restrikčních fragmentů (RFLP), 10 polymorfismus délky amplifikovaných fragmentů (AFLP), opakování jednoduchých sekvencí (SSR), jedno-nukleotidový polymorfismus (SNP) a DArT markery (Gupta et al., 2008).

Právě SNP markery se v posledních studiích ukazují jako velmi vhodné pro genomickou selekci hospodářských zvířat a rostlin, jelikož vykazují vysokou účinnost v korelacích s kýženými fenotypickými znaky, dále se vyznačují vysokou reprodukovatelností, stupněm polymorfismu, 15 abundancí v genomu ad. v porovnání s metodami alternativními (Nadeem et al., 2017).

Užitný vzor si klade za úkol navrhnout nástroj pro přípravu populačního čipu na bázi sond pro identifikaci konkrétních markerů jednonukleotidových polymorfismů (SNP), které odpovídají 20 variacím v sekvenci DNA, k nimž dochází při záměně jednoho nukleotidu, tj. adeninu (A), thyminu (T), cytosinu (C) nebo guaninu (G) v sekvenci genomu.

#### Podstata technického řešení

25

Podstata technického řešení spočívá v testovací sadě k přípravě populačního čipu pro efektivní genotypování a selekci jetele lučního prostřednictvím genů asociovaných s vyšší biologickou fixací dusíku *T. pratense* odrůdy Tatra. Testovací sada je složena z primeru 1, tj. komplementární cílové sondy k vybraným genům, a z primeru 2 pro přípravu amplikonových sond s linkerem. Přitom 30 primery jsou tvořeny sekvencemi nukleotidů uvedenými v tab. 1.

Tabulka 1

| č. | lokus    | pozice | Primer 1                       |
|----|----------|--------|--------------------------------|
| 1  | Tp_9375  | 133    | ATTTTCGGCGTAGGACA              |
| 2  | Tp_13036 | 2625   | ATCTTTCCTTGATGTAACCTGT         |
| 3  | Tp_7124  | 4424   | CATGTTGTGAGAGATGTTGA           |
| 4  | Tp_86    | 3276   | ACATGTTATATCAGTGTAATTTGA       |
| 5  | Tp_2269  | 7537   | TCAAATAATTGTTTCCTTTCCTTTGA     |
| 6  | Tp_31403 | 2100   | AATTGGTTTGTTTAGATTAGCG         |
| 7  | Tp_8398  | 805    | TTGATGCAAAAGGGATGGAGA          |
| č. | lokus    | pozice | Primer 2                       |
| 1  | Tp_9375  | 133    | [AmC6]TCCACAAAAGGAACTTATAATAA  |
| 2  | Tp_13036 | 2625   | [AmC6]TGAAGAGCTTATTTTCAGATAATT |
| 3  | Tp_7124  | 4424   | [AmC6]CTTCCGGTTGAGATTTCCA      |
| 4  | Tp_86    | 3276   | [AmC6]CAAGATTGAAATTACAGTTTGT   |
| 5  | Tp_2269  | 7537   | [AmC6]TAGAGAAAAGGCACAAGCG      |
| 6  | Tp_31403 | 2100   | [AmC6]AATCATGGTAGCACAGC        |
| 7  | Tp_8398  | 805    | [AmC6]CCGTCACAGCAACGTGATAC     |

- 5 [AmC6] může být libovolný linker k ukotvení vzniklého amplikonu na substrát k přípravě popsaného populačního čipu. Přitom linker zahrnuje chemické skupiny schopné vytvořit vazbu na uhlíkatý řetězec DNA amplikonu prostřednictvím funkční skupiny, jako např. aldehydická, amino, epoxy a další, případně pomocí systémů založených na vazbách s biotinem a streptavidinem a odvozených systémů.
- 10 Geny jedinců *T. pratense* vybrané pro selekci na bázi SNP markerů byly přitom paralelním výzkumem asociovány a lokalizovány u specifických genů, uvedených v tab. 2, jako významné pro biologickou fixaci dusíku a procesy nodulace. Výběr genů byl proveden na základě experimentálních výsledků z měření fixace dusíku, případně na základě jejich specifické exprese a zahrnuje následující rodiny genů: transkripční faktory, leghemoglobiny, receptorové kinázy
- 15 a cytokininové receptory.

Tabulka 1:

| označení genu | název genu                          | funkce genu  |
|---------------|-------------------------------------|--|
| Tp_9375       | <i>ERF required for nodulation1</i> | transkripční faktor, kontrola nodulace, regulace růstu infekčních vláken |
| Tp_13036      | <i>PEN3-like</i>                    | rezistence vůči chorobám, specificky exprimovaný v hlízkách              |
| Tp_7124       | <i>SUNN</i>                         | kináza, kontrola počtu hlízek  |
| Tp_86         | <i>PHO2-like</i>                    | regulace akumulace fosforu v hlízkách                                    |
| Tp_2269       | <i>NFP</i>                          | receptor Nod faktorů   |
| Tp_31403      | <i>Ca/CaM-dependent kinase</i>      | přenos signálu Nod faktoru, nodulace                                     |
| Tp_8398       | <i>RDNI</i>                         | regulace počtu hlízek  |

Objasnění výkresů

5 Technické řešení bude dále objasněno pomocí výkresů, na nichž obr. 1 ukazuje počáteční řádky popsaného souboru biologických dat vytvořeného ve formátu GAL (GenePix Array List), obr. 2 představuje vizualizaci celkového náhledu 5 podpolí o 8 blocích z tiskového software a obr. 3 náhled skenu populačního čipu exportovaného do formátu \*.jpn.

Příklady uskutečnění technického řešení

10

Prostřednictvím testovací sady byl navržen a vytvořen nový typ populačního čipu s kvantitativním spektroskopickým stanovením fluorescence. Principem je interakce krátkých vybraných úseků DNA, které jsou imobilizovány na aktivované matrici, s jejich komplementárními fluorescenčně značenými sondami a které detekují vybrané SNP. Populační čip se skládá z mikroskopického skla  
15 (25x75 mm) chemicky modifikovaného pro imobilizaci vybraných krátkých úseků sekvencí genů (amplikonových sond relevantních pro fixaci dusíku), získaných PCR amplifikací tzv. linkerem (spojovací molekulou) k vazbě na modifikovaný substrát z každého vzorku, a z kontrolních sekvencí.

20

Populační čip byl vytvořen použitím tiskového robotu NanoPrint™ (Arrayit) s použitím tiskové hlavy a pinů SMP3 (Telechem). Geometrie rozmístění spotů na čipu může být značně rozdílná v závislosti na počtu zvolených cílových sekvencí a vzorků. Pro výtisk PCR sond bylo použito rozmístění do 5 sub-polí 450 μm od sebe vzdálených po 8 blocích, přičemž v každém bloku jsou spoty 10 x 11 s mezerou 420 x 400 μm. Každý loci a kontroly jsou vytištěny ve 4 opakováních  
25 v podobě spotu o průměru přibližně 100 μm (u velikosti spotu je možná variabilita). Výše zmíněné parametry populačního čipu a dále informace k identifikaci pozic jednotlivých vzorků na substrátu jsou obsaženy v tzv. GAL souboru.

30

Soubor specifických primerů je určen pro přípravu populačních amplikonových sond obsahujících SNP polymorfismy přibližně uprostřed každé uvedené sekvence. Tyto amplikonové sondy jsou následně roboticky nanášeny na substrát, kde dojde k jejich ukotvení prostřednictvím výše zmíněného linkeru, čímž je připraven populační čip. Hybridizace na čipu je provedena za vhodných reakčních podmínek (promývání pufrů a při optimalizovaném teplotním režimu) s fluorescenčně značenými syntetizovanými cílovými sondami se specifickou sekvencí pro dané SNP.  
35

35

Všechny sekvence těchto sond byly odvozeny z referenční sekvence *T. pratense* odrůdy Tatra, přičemž alelové varianty jsou odlišeny různým fluoroforem (zpravidla červeným a zeleným dle nastavení kanálů v dostupném fluorescenčním skeneru). Následné vyhodnocení fluorescenčních signálů umožňuje relativně jednoduché a rychlé genotypování vybraných SNP markerů v celé populaci.  
40

40

K robotické přípravě čipů se využívají kontaktní a bezkontaktní metody depozice buď plně funkčních molekul, nebo syntetizovaných in-situ na povrchu substrátu (nejčastěji mikroskopická sklíčka s chemicky modifikovaným povrchem) zahrnující fotolitografickou syntézu (Lipshutz et al. 1999, Singh-Gasson et al. 1999) nebo syntézu pomocí technologie piezoelektrického tisku (Hughes et al. 2001). Základní koncept populačních čipů spočívá v tom, že každý spot poskytuje jedinečný signál v důsledku unikátní hybridizace cílových molekul se sondami a zároveň populační čip poskytuje mnohačetnou paralelní detekci cílených markerů.  
45

45

50

Konkrétní rozložení vzorků na designovaném populačním čipu je řešeno pomocí GAL souboru sloužícího k identifikaci jednotlivých pozic vzorků na substrátu. Příklad takového technického řešení je vidět na obr. 1, jenž ukazuje počáteční řádky popsaného GAL souboru. Obr. 2 dále představuje vizualizaci celkového náhledu 5 podpolí o 8 blocích z tiskového software.

Po přípravě je populační čip podroben mikroskopické kontrole kvality a následně při zvýšené teplotě hybridizován se syntetickými sondami (2 až 2,5  $\mu\text{l}/\text{cm}^2$ ), značenými flouorofory, komplementárními k hledaným loci (cílové sondy; fluorofor Cy3 pro alelu divokého typu *wt* (z angl. wild type) a Cy5 pro mutantní alelu *m* (z angl. mutant). Vyhodnocení bylo provedeno na fluorescenčním skeneru GenePix® 4000B (Molecular Devices) s detekčními kanály  $\lambda = 532 \text{ nm}$  a  $\lambda = 635 \text{ nm}$ . Náhled skenu populačního čipu lze exportovat např. do formátu \*.jpg, jak je ukázáno na obr. 3.

DNA sekvence (amplikonové sondy) pro výtisk byly připraveny v rozmezí koncentrací 0,1 až 1  $\mu\text{g}/\mu\text{l}$ , kontroly 1 až 50  $\mu\text{M}$ . Z intenzity signálu byly určeny koeficienty pro genotypy daných loci a vzorku odpovídající homo/heterozygotnosti následujícím způsobem:

Ze získaných dat byly odfiltrovány hodnoty nesplňující kritérium  $\text{SNR} > 3$  u vlnových délek 532 nm a 635 nm (z angl. Signal to Noise Ratio, kde parametr udává, zda signál ve spotu je dostatečně silný oproti pozadí) a tzv. “flagů” (funkce sloužící k filtraci a třídění dle kvality spotů, např.: dobrý, špatný, chybný, poškozen). Následně byly odfiltrovány monoplíkáty a provedena kontrola replikátů ( $n > 2$ ). Vzorek per loci byl odstraněn, pokud nebyly přítomny alespoň 2 repetice. Proběhla kontrola distribuce kontrol dle genotypu/lokusu a následně normalizace per loci tak, aby se poměr intenzit u kontrol pro heterozygota rovnal 1. Kontroly byly po normalizaci podrobeny druhé kontrole distribuce, v případě nutnosti filtrovány a po filtraci byla data znovu normalizována. Replikáty byly sumarizovány na medián a transformovány  $\log 2$ . Z kontrol byl vytvořen lineární SVM model, dle kterého došlo k následné klasifikaci dat. SVM model (z angl. Support Vector Machines), známý též jako metoda podpůrných vektorů, je metodou používanou ke klasifikaci a pro regresní analýzu. Finální soubor s daty pak obsahuje:  $\log 2$  normalizované intenzity, ID, pravděpodobnost příslušnosti k jednotlivým genotypům a výsledný přiřazený genotyp, tzn. ten s největší pravděpodobností.

Populační čip je určen k efektivní genotypizaci populací *T. pratense* odrůdy Tatra pro alely s asociací vyšší aktivity fixace vzdušného dusíku. Pracovní postup zahrnuje: izolaci DNA ze vzorků rostlin, PCR amplifikaci s modifikací linkerem pro ukotvení na substrát populačního čipu, purifikaci PCR produktu, výtisk amplikonů na substrát, jejich chemickou imobilizaci, hybridizaci s cílenými syntetickými sondami značenými flouorofory a finální bioinformatické vyhodnocení. Populační čip byl navržen tak, aby prostřednictvím identifikace vybraných alelových variant SNP umožnil efektivní selekci v rámci šlechtitelských populací (řádově stovky až desetitisíce jedinců), což šlechtitelé využijí pro efektivní predikci žádaného fenotypového znaku, tj. vyšší biologickou fixaci dusíku v rostlinách.

Inovativnost řešení spočívá v současné mnohačetné analýze vybraných genetických znaků, jednonukleotidových polymorfismů (SNP) se vztahem k biologické fixaci dusíku, analyzovaných současně na 1 populačním čipu. Tato technologie je cílena na uplatnění v oblasti šlechtění a selekce jetele lučního, konkrétně na selekci jedinců s genetickými znaky nesoucími vyšší schopnost rostliny fixovat vzdušný dusík a tím ekologicky pomáhat k obohacení půdy dusíkem.

## NÁROKY NA OCHRANU

1. Testovací sada k přípravě populačního čipu pro efektivní genotypování a selekci jetele lučního prostřednictvím genů asociovaných s vyšší biologickou fixací dusíku *T. pratense* odrůdy Tatra, která je složená z primeru 1, tj. komplementární cílové sondy k vybraným genům, a z primeru 2 pro přípravu amplikonových sond s linkerem, **vyznačující se tím**, že primery jsou tvořeny následujícími sekvencemi nukleotidů:

| č. | lokus    | pozice | Primer 1                       |
|----|----------|--------|--------------------------------|
| 1  | Tp_9375  | 133    | ATTTTCGGCGTAGGACA              |
| 2  | Tp_13036 | 2625   | ATCTTTCCTTGATGTAACCTGT         |
| 3  | Tp_7124  | 4424   | CATGTTGTGAGAGATGTTGA           |
| 4  | Tp_86    | 3276   | ACATGTTATATCAGTGTAATTTGA       |
| 5  | Tp_2269  | 7537   | TCAAATAATTGTTTCCTTTCCTTTGA     |
| 6  | Tp_31403 | 2100   | AATTGGTTTGTGTTAGATTAGCG        |
| 7  | Tp_8398  | 805    | TTGATGCAAAAGGGATGGAGA          |
| č. | lokus    | pozice | Primer 2                       |
| 1  | Tp_9375  | 133    | [AmC6]TCCACAAAAGGAACTTATAATAA  |
| 2  | Tp_13036 | 2625   | [AmC6]TGAAGAGCTTATTTTCAGATAATT |
| 3  | Tp_7124  | 4424   | [AmC6]CTTCCGGTTGAGATTTCCA      |
| 4  | Tp_86    | 3276   | [AmC6]CAAGATTGAAATTACAGTTTGT   |
| 5  | Tp_2269  | 7537   | [AmC6]TAGAGAAAAGGCACAAGCG      |
| 6  | Tp_31403 | 2100   | [AmC6]AATCATGGTAGCACAGC        |
| 7  | Tp_8398  | 805    | [AmC6]CCGTACAGCAACGTGATAC      |

10

příčemž [AmC6] může být libovolný linker k ukotvení vzniklého amplikonu na substrát k přípravě popsaného populačního čipu a příčemž linker zahrnuje chemické skupiny schopné vytvořit vazbu na uhlíkatý řetězec DNA amplikonu prostřednictvím funkční skupiny, jako např. aldehydická, amino, epoxy a další, případně pomocí systémů založených na vazbách s biotinem a streptavidinem a odvozených systémů.

15

3 výkresy

```

ATF      1.0
#6       5
"Type=GenePix ArrayList V1.0"
"BlockCount=40"
"BlockType=0"
"Block1= 3000, 3000, 100, 11, 400, 10, 420"
"Block2= 7500, 3000, 100, 11, 400, 10, 420"
"Block3= 12000, 3000, 100, 11, 400, 10, 420"
"Block4= 16500, 3000, 100, 11, 400, 10, 420"
"Block5= 3000, 7500, 100, 11, 400, 10, 420"
"Block6= 7500, 7500, 100, 11, 400, 10, 420"
"Block7= 12000, 7500, 100, 11, 400, 10, 420"
"Block8= 16500, 7500, 100, 11, 400, 10, 420"
"Block9= 3000, 12280, 100, 11, 400, 10, 420"
"Block10= 7500, 12280, 100, 11, 400, 10, 420"
"Block11= 12000, 12280, 100, 11, 400, 10, 420"
"Block12= 16500, 12280, 100, 11, 400, 10, 420"
"Block13= 3000, 16780, 100, 11, 400, 10, 420"
"Block14= 7500, 16780, 100, 11, 400, 10, 420"
"Block15= 12000, 16780, 100, 11, 400, 10, 420"
"Block16= 16500, 16780, 100, 11, 400, 10, 420"
"Block17= 3000, 21560, 100, 11, 400, 10, 420"
"Block18= 7500, 21560, 100, 11, 400, 10, 420"
"Block19= 12000, 21560, 100, 11, 400, 10, 420"
"Block20= 16500, 21560, 100, 11, 400, 10, 420"
"Block21= 3000, 26060, 100, 11, 400, 10, 420"
"Block22= 7500, 26060, 100, 11, 400, 10, 420"
"Block23= 12000, 26060, 100, 11, 400, 10, 420"
"Block24= 16500, 26060, 100, 11, 400, 10, 420"
"Block25= 3000, 30839, 100, 11, 400, 10, 420"
"Block26= 7500, 30839, 100, 11, 400, 10, 420"
"Block27= 12000, 30839, 100, 11, 400, 10, 420"
"Block28= 16500, 30839, 100, 11, 400, 10, 420"
"Block29= 3000, 35339, 100, 11, 400, 10, 420"
"Block30= 7500, 35339, 100, 11, 400, 10, 420"
"Block31= 12000, 35339, 100, 11, 400, 10, 420"
"Block32= 16500, 35339, 100, 11, 400, 10, 420"
"Block33= 3000, 40120, 100, 11, 400, 10, 420"
"Block34= 7500, 40120, 100, 11, 400, 10, 420"
"Block35= 12000, 40120, 100, 11, 400, 10, 420"
"Block36= 16500, 40120, 100, 11, 400, 10, 420"
"Block37= 3000, 44620, 100, 11, 400, 10, 420"
"Block38= 7500, 44620, 100, 11, 400, 10, 420"
"Block39= 12000, 44620, 100, 11, 400, 10, 420"
"Block40= 16500, 44620, 100, 11, 400, 10, 420"
"Supplier=Essenceline"
"ArrayerSoftwareName=Microarray Manager"
"ArrayerSoftwareVersion=3, 0, 0, 32"
"Block" "Column" "Row" "ID" "Name"
1      1      1      K35/8_C3      K35/8_C3
2      1      1      K35/8_C3      K35/8_C3
3      1      1      K35/8_C3      K35/8_C3
    
```

Obr. 1







Obr. 3