



Biofyzikální ústav Akademie věd České republiky

RNDr. Aleš Kovářík, CSc., Královopolská 135, 612 65 Brno,

Czech Republic, tel.: 541 517 178, e-mail: kovarik@ibp.cz

Oponentský posudek na disertační práci "Konstrukce fyzické mapy krátkého ramene chromozómu 7D pšenice a její využití pro poziční klonování "

Autorka: Mgr. Helena Staňková

Pracoviště: Univerzita Palackého v Olomouci a Ústav experimentální botaniky AVČR

Pšenice patří s rýží a kukuřicí k nejdůležitějším plodinám a je základním zdrojem obživy pro třetinu lidstva. Statistiky udávají, že každoročně se jí sklídí asi 700 milionů tun. V posledních desetiletích se jí však sklídilo méně, než lidstvo spotřebovalo, a logickým důsledkem je pokles světových zásob. S nárůstem lidské populace se však poptávka po pšenici bude v budoucnu neustále zvyšovat a zajistit požadovanou produkci pšenice v roce 2050 bude možné jen tehdy, pokud budou výnosy nových odrůd stoupat o 2 % ročně. Takového růstu lze však dosáhnout jen intenzivním šlechtěním, které využívá metody molekulární biologie a genomiky a také biotechnologické metody, včetně genetické transformace a editace dědičné informace. Rozsáhlejší zapojení těchto metod do šlechtitelských programů a plné využití jejich potenciálu však nebude uskutečnitelné bez podrobné znalosti dědičné informace (genomu) pšenice. K tomuto cíli směřuje i téma předložené disertace, která je členěna do dvou organicky na sebe navazujících částí.

První část práce je věnovaná pozičnímu klonování genu(ů) v lokusu *Dn2401*, jenž podmiňuje rezistenci plodiny k mšici zhoubné. S cílem zahustit genetickou mapu byl navržen postup pro cílený vývoj markerů v prostředí hexaploidního pšeničného genomu. Za použití této metody bylo vyvinuto několik nových markerů v těsné vazbě na lokus, pomocí kterých byl identifikován kontig fyzické mapy, jenž překrývá oblast zodpovídající za rezistenci. Sekvenování a následná anotace sekvencí klonů ze získaného kontigu odhalila přítomnost několika kandidátních genů. Ačkoliv gen resistance izolován nebyl, jeho lokalizace se zúžila na poměrně krátkou genetickou vzdálenost odpovídající několika stovkám kb. Je tudíž pravděpodobné, že se tak v blízké budoucnosti stane.

V druhé části práce byla zkonstruována fyzická kontigová mapa ramene 7DS. Ta posloužila k výběru tzv. *minimal tilling path* (MTP), tedy minimální sestavy klonů z knihovny dlouhých

inzerťů, která reprezentuje celé chromozómové rameno. Klony z MTP byly následně sekvenovány a získaná data se stala základem pro sestavení referenční sekvence ramene 7DS. Za účelem ověření správnosti fyzické mapy a tedy i výsledné chromozómové sekvence bylo provedeno srovnání s genomovou mapou vytvořenou technologií BioNano na platformě Irys. Toto srovnání prokázalo, že optické mapování může posloužit jako užitečný nástroj pro ověřování a zdokonalení fyzické mapy a také se může uplatnit při analýze komplikovaných částí genomu, jako jsou tandemové repetice.

Všechny výsledky presentované v dizertaci prošly náročným recenzním řízením, takže oponentovi nenáleží podrobovat kritice již uveřejněné výsledky. Autorčin příspěvek je zřejmý – prvoautorství na dvou kvalitních vysoce impaktovaných publikacích. Přesto si z vlastního zájmu dovolím položit několik dotazů:

1. V publikaci byla metodou optického mapování odhalena nová oblast chromosomu, která dosud nebyla pokryta sekvenovanými BACovými klony. Tento výsledek je velmi zajímavý ukazující na užitečnost této relativně nové metody skýtající možnosti jejího širšího uplatnění v genomických projektech. Nově objevený úsek obsahuje opakující se segment se štěpnými místy pro nikující restrikčním enzym. Hypotéza, že se jedná o satelitní sekvenci je však značně spekulativní a není podložena příslušnými experimenty. Je stejně docela možné, že se jedná o úseky kódující 35S ribosomální RNA (rDNA) vzhledem k tomu, že perioda opakování odpovídá délce monomerní jednotky rDNA u pšenice (cca 10 kb, Gerlach a Bedbrook, 1979) a také proto, že se jeden z 35S rDNA lokusů pšenice na tomto chromosomu vyskytuje (Mukai a spol 1991). Zajímalo by mne, zda by mohla autorka nastínit experimentální postup, který by vedl k bližší charakterizaci zmíněného úseku chromosomu. Že se jedná o rDNA by možná šlo vyloučit analýzou restrikčních míst monomerní jednotky rDNA. Pokud se cílové místo pro použitý restrikční enzym *Nt.BspQI* v sekvenci nevyskytuje nebo se naopak vyskytuje mnohokrát (alespoň 2x), patrně se u mapy č. 350 (Obr. 1) nejedná o klast rDNA. Z evolučního hlediska by bylo rovněž zajímavé zjistit úroveň kolinearity mezi chromosomem 7DS u pšenice a chromosomy jejího diploidního předka.

2. Dizertační práce je do značné míry orientovaná na aplikaci metod pro výzkum velkých genomů. Například jedním z hlavních cílů byla validace fyzické mapy krátkého ramene chromosomu pšenice optickým mapováním. S výjimkou literárního úvodu (krátké odstavce na str. 22 a 40) postrádám však detailnější popis této relativně nové metody. Obvyklé části "Materiál a Metody" nejsou totiž v dizertaci uvedeny vůbec. Rád bych proto požádal autorku, aby mohla v krátkosti během své ústní prezentace shrnout dosavadní poznatky o Bionano technologii. Poukazuji zde na literaturu například: *Hastie et al.:Rapid Genome Mapping in*

Nanochannel Arrays for Highly Complete and Accurate De Novo Sequence Assembly of the Complex Aegilops tauschii Genome. PlosOne February 6, 2013, DOI:

10.1371/journal.pone.0055864. V polemice na téma BioNano by mne zajímaly technické detaily jako je: (i) počet restrikčních enzymů, které se uplatňují při fluorescenčním značení analyzovaných molekul, (ii) počet využitelných fluorochromů pro značení. (iii) minimální množství výchozího vzorku. (iv) rozlišovací schopnost metody. Jinými slovy, jaká je nejkratší vzdálenost mezi jednotlivými restrikčními místy, kterou lze ještě opticky detekovat?

3. Práce je po metodické stránce extrémně bohatá ukazující na technickou zdatnost doktorandky. Zejména vyzdvihují originální přístup pro izolaci vysokomolekulární DNA metodou sortrovaných chromosomů na průtokovém cytometru. Tato metoda se osvědčila při přípravě DNA pro konstrukci BAC knihoven a molekul pro optické mapování. Vzhledem k tomu, že na publikacích se podíleli členové domácích i zahraničních výzkumných týmů, prosím autorku, aby ve své ústní presentaci uvedla, které metody vlastnoručně prováděla a u kterých by mohla poskytnout odbornou expertízu případným zájemcům.

4. Strana 27 el. textu, kapitola 2.2.3.1. Věta *"Sequence analysis showed a highly dynamic structure of the wheat genome with significant loss of genes (between 10,000 and 16,000) compared to..."* je v rozporu s větou na straně 14 *"Comparison of the hexaploid bread wheat gene sequences with gene repertoires from its closest extant relatives, which donated the A, B and D genomes showed limited gene loss during the evolution of the hexaploid wheat genome and frequent gene duplications after these genomes came together (Eversole et al. 2014)"*. Dle mého názoru je ztráta 10-16 tisíc genů významná a nelze ji považovat za "limited" uvědomíme-li si, že jeden genom pšenice kóduje kolem 90 tisíc genů. Navíc několik studií prokázalo změny v nekodujících repetitivních oblastech, kde se jednalo většinou o rozsáhlé Mbázové delece.

Rovněž na úrovni exprese nesouhlasím s tvrzením *"Gene expression patterns revealed that none of the subgenomes dominated gene expression"*. Mnoho izolovaných genů totiž genom-specifické umlčování vykazuje. Známým příkladem je dominance lokusu rDNA na chromosomech 1B a v menším míře 6B oproti umlčeným lokusům na chromosomech 5D, 1A, a již zmiňovaném chromosomu 7D.

5. Z textu není příliš zřejmé, zda pro optické mapování byla použita stejná linie pšenice jako pro konstrukci BAC knihovny. Prosím o vysvětlení. Možná, že by bylo také zajímavé analyzovat optickým mapováním linii obsahující lokus resistance *Dn2401* eventuálně další intenzivně studované kultivary pšenice a diploidní rodiče a sledovat genetickou variabilitu úseku odpovídajícího mapě 350.

6. Terminologie je v některých případech ne zcela přesná. Například v případě studovaného *Dn2401* se jedná o lokus a nikoliv gen (str. 77 a jinde). Gen resistance teprve bude izolován.

Drobné připomínky

Strana 23 el. textu, kapitola 2.2.3. "cannot be obtain through" má být trpný rod – "obtained"

Získané sekvence by měly být submitovány do veřejné database

Strana 13.: "Each hybridization was followed by a polyploidization" Jaké jsou důkazy pro tuto hypotézu? Alternativní cesta vzniku allopolyploida je prostřednictvím hybridizace neredukovaných gamet. V tomto případě vzniká polyploidní jádro ihned po mezidruhovém křížení.

Závěr

Disertační práce Mgr. Staňkové obsahuje řadu kvalitních originálních poznatků, které byly nebo budou v budoucnu publikovány. Výsledky významně přispěly do mozaiky našich poznatků o genomu pšenice a mohou nalézt uplatnění ve šlechtitelství a biotechnologiích. Kvalitou práce převyšuje celostátní průměr v daném oboru. Je to dáno v neposlední řadě i kvalitním vedením a úrovní pracoviště v AVČR.

Práci doporučuji k obhajobě

V Brně, dne 30. listopadu, 2015



Aleš Kovařík

BIOFYZIKÁLNÍ ÚSTAV AV ČR, v.v.i.
Královopolská 135, 612 65 BRNO
IČ: 68081707, DIČ: CZ68081707



Přírodovědecká
fakulta

Posudek disertační práce Mgr. Heleny Staňkové

Studijní program P1527 Botanika

Opponent statement of the dissertation titled:

Construction of physical map of 7DS wheat chromosome arm and its use for positional cloning.

Thesis aims to:

- 1) Construction of a physical contig map for the short arm of the wheat chromosome 7D, select and sequence minimal tilling path (MTP), and finally validate the physical map by an alternative mapping technology.
- 2) Construction of a high-density genetic map in Dn2401 gene region The second aim of the thesis was to construct a high-density genetic map in the region of an aphid resistance gene - Dn2401. This key step in the positional cloning process should enable to span the region by a physical map contig and potentially identify candidate gene(s).

The introductory parts cover sufficiently comprehensively wheat domestication and genomics issues. Due to polyploid hybrid origin of bread wheat it focuses on reductions of genome complexity, particularly by approach developed by prof. J. Doležel team, e.g. chromosomal genomics based on flow sorting of individual chromosomes or even chromosome arms. Excellently is provided the up to date overview of plant genome sequencing technologies currently used such as whole genome shotgun and clone-by-clone approaches. As further work is on physical mapping and high-density genetic mapping these are explained in more details further. Since project within which this thesis was carried out aims to isolate the gene responsible for resistance to Russian wheat aphid (*Diuraphis noxia*), some relevant information to this pathogen are provided too. All these together cover 50 pages followed by comprehensive list of references indicating good knowledge of Ph. D. candidate.

Results are entirely presented in form of published impacted papers with first authorship as required for granting Ph.D. These were rigorously reviewed by experts in the field. On the other hand these very compacted data do not allow too much space for discussion as only highest quality outcome is shown. This is pity in some sense. Author's contribution is clearly stated and it is obvious that this is multi-team task, as common in today's science. In my view would be useful to provide also separate author's more personal conclusions and perspectives on the end of the thesis. Mentioning also very likely large body of yet unpublished data and pitfall met during the work. Thesis offers this opportunity rather than peer reviewed papers. Not entirely nice is structuring, using far too long sub-numbering (such as 2.2.3.2.2.4) simplification would be possible and useful.

One accepted for publication at Plant Biotechnology Journal (IF: 5.752) *BioNano genome mapping of individual chromosomes supports physical mapping and sequence assembly in complex plant genomes*. Staňková H et al.

Here authors used short arm of chromosome 7D (7DS) as a model to demonstrate for the first time that it is possible to couple chromosome flow sorting with genome mapping in nanochannel arrays and create a de novo genome map of a wheat chromosome. They constructed a high resolution chromosome map composed of 371 contigs with an N50 of 1.3 Mb. BioNano mapping approach facilitated chromosome-

scale analysis of repetitive sequences and revealed a ~800-kb array of tandem repeats intractable to current DNA sequencing technologies.

Second paper was already published at Theoretical and Applied Genetics, 128(7):1373-1383, 2015. (IF: 3.790) *Chromosomal genomics facilitates fine mapping of a Russian wheat aphid resistance gene*. Staňková H et al.

This paper focuses to clone a Russian wheat aphid resistance gene Dn2401 located on wheat chromosome arm 7DS using information gathered by Genome Zipper, a synteny-based tool combining sequence data of rice, *Brachypodium*, sorghum and barley, and taking advantage of a high-density linkage map of *Aegilops tauschii*. All these added in an efficient and cost effective manner 11 new gene-associated markers in the Dn2401 region and narrow down the target interval to just 0.83 cM. Screening 7DS-specific BAC library with the flanking markers revealed a contig of four BAC clones that span the Dn2401 region in wheat cultivar 'Chinese Spring'. Furthermore author has presented her results in form of 4 posters at international conferences, which are added to the thesis.

In summary, the thesis is very well written and present comprehensive view on wheat chromosome genomics. Based on the novelty of research reported and moving the science forward of large complex genomes in general and wheat specifically with published and publishable research, in my opinion, this presentation meets the standard for the completion of the Doctor of Philosophy for the dissertation portion degree.

In Olomouc 26 November 2015



Dr. Petr Smýkal,
Department of Botany, Faculty
of Sciences, Palacký University
in Olomouc

Oponentský posudek na Disertační práci Heleny Staňkové "Konstrukce fyzické mapy krátkého ramene chromozómu 7D pšenice a její využití pro poziční klonování"

Disertační práce Heleny Staňkové "Konstrukce fyzické mapy krátkého ramene chromozómu 7D pšenice a její využití pro poziční klonování" byla vypracována Katedře botaniky Přírodovědecké fakulty University Palackého a na Ústavu experimentální botaniky AV ČR, Centru regionu Haná pro biotechnologický a zemědělský výzkum. Práci vedla ing. Hana Šimková CSc. Disertační teze jsou sepsány v anglickém jazyce.

Práce se skládá ze dvou základních částí. První z nich je tvořena detailním, přehledně členěným úvodem (60 stran). Druhou část představují vlastní výsledky v podobě dvou publikací a čtyř plakátových sdělení. Helena Staňková je první autorkou obou článků publikovaných v prestižních časopisech Plant Biotechnology Journal a Theoretical and Applied Genetics.

Doktorandský výzkum se týká dvou témat – fyzického mapování a skládání sekvence krátkého ramene chromozomu pšenice 7D a využití chromozomových map a sekvenční informace pro poziční klonování genu *Dn2401*, který se nalézá na 7DS a podmiňuje rezistenci k mšici zhoubné. Pšeničný genom je značně veliký (asi 17 Gb), jeho podstatnou část tvoří repetitivní sekvence. Skládá se ze tří subgenomů, odvozených od diploidních předků – *Triticum urartu* (A), *Aegilops speltoides* (B) a *Aegilops tauschii* (D). Tato hexaploidní konfigurace znamená, že každý gen se může vyskytovat až v šesti alelách. Nalezení kompletní sekvence pšeničného genomu prostě chybí proto představuje tvrdý oříšek na samých hranicích současného vědeckého poznání. Přitom se rovněž jedná o velmi praktický úkol, protože pšenice je doslova naším denním chlebem. Dosud pracovali šlechtitelé s náhodnými mutacemi a znaky, které vybírali a kombinovali, avšak nedokázali je cíleně indukovat. Kompletní genomická sekvence umožní předem navrhnout genetické složení nové odrůdy, projektovat genom „na míru“.

Extrémní velikost pšeničného genomu vyžaduje vypracování nových kombinovaných metod sekvenování, které mohou být následně použity pro genomickou analýzu ostatních velkých polyploidních rostlinných genomů. Laboratoř molekulární cytogenetiky a cytometrie Ústavu experimentální botaniky AV ČR a Centra regionu Haná pro biotechnologický a zemědělský výzkum, kde byla disertace vypracována, má dlouholeté zkušenosti s tříděním chromozomů. Tento postup byl použit pro získání suspenze krátkého ramene chromozomu 7D z ditelosomické linie pšenice kultivaru „Chinese Spring“. Telocentrické chromozomy odpovídající dvěma ramenům 7D chromozomu jsou kratší než zbývající chromozomy a mohou být tříděny na základě odlišné velikosti. Rozdělení celkového genomu na jednotlivé chromozomy či jejich ramena umožňuje sekvenovat genom po částech, čímž se podstatně redukuje komplexita.

Chromozom 7DS je dlouhý asi 381 Mb, DNA k sekvenování je tedy zhruba 50x kratší. Ale i 381 Mb je příliš mnoho pro přímé shotgun (brokovnicové) sekvenování pomocí např. Illumina metody, která produkuje reads o délce několika set bp. Vždyť délka tohoto ramene je srovnatelná s celým genomem modelové rostliny *Arabidopsis thaliana*. Proto i v případě jednotlivých chromozomů postupuje sekvenování po menších částech, jsou konstruovány BAC knihovny s inzerty o velikosti desítek až stovek bp, které jsou sekvenovány např. pomocí Illumina. K sestavení celkové sekvence chromozomového ramene je pak nutné správně

seřadit jednotlivé BAC sekvence i jejich kontigy, ukotvit je k mapě chromozomu. Toto přiřazení je prováděno na základě známých genetických markerů tvořících genetickou mapu. Pozice markeru v mapě chromozomu je srovnána s jeho výskytem v daném BAC inzertu či kontigu.

Sestavení genetické mapy je zdlouhave a obtížné, navíc závisí na rekombinaci, jejíž frekvence je nízká v některých oblastech genomu. Alternativou se proto stává optické mapování, které produkuje fyzické mapy genomu s body odpovídajícími místům rozpoznávaným „nickasami“. Tyto endonukleázami provádějí místně specifické jednořetězcové zlomy vzdálené od sebe desítky až stovky kb, a proto vhodné k analýze velmi dlouhých vláken DNA. Použití nanokanálek pro rovnoměrné natažení vláken DNA představuje současné zdokonalení metody optického mapování.

Mapování v nanokanálcích, tzv. BioNano mapování, použila Helena Staňková k sestavení náčrtu sekvence krátkého ramene pšeničného chromozomu 7D. Kombinace třídění chromozomů pomocí průtokové cytometrie a BioNano genomického mapování je novým metodickým přístupem vypracovaným v Laboratoři molekulární cytogenetiky a cytometrie. Ukázal se být velmi vhodným k *de novo* sekvenování velkých polyploidních genomů. Jeho aplikaci se zabývá publikace – Staňková et al.: BioNano-genome mapping of individual chromosomes supports physical mapping and sequence assembly in complex plant genomes - přijatá v Plant Biotechnology Journal. Autorka společně s 12 spoluautory sestavila 371 genomických map pokrývajících souhrnně 350 Mb a představujících tedy 92% délky 7DS. Průměrná délka genomické mapy byla 0.9 Mb, N50 rovno 1.3 Mb. V průběhu mapování byla nalezena tandemová repetice s délkou jednotky asi 9.25 kb, která zahrnovala celkem 1 Mb. Takovéto repetice nelze přesně sestavit pomocí sekvenování, protože se obtížně klonují a krátké reads neumožňují stanovit přesný počet jednotek. BioNano mapování ji velmi dobře popsalo, jak názorně ukazuje Obr. 1. Toto mapování rovněž pomohlo odhalit nesprávné složení některých kontigů sestavených z BAC inzertů.

V další části práce se kandidátka zabývala podrobným genetickým mapováním části 7DS chromozomu, která je zodpovědná za rezistenci k mšici zhoubné (*Dn2401*). Využila syntenie se známými genomy *Aegilops tauschii* a *Brachypodium* a také dříve sestavených genomických map 7DS. Nalezla 11 nových genomově specifických markerů pro oblast *Dn2401*. Výraznou úlohu hrálo i zde třídění chromozomů, které byly např. použity k ověření genomické specifity nových markerů. Tuto práci publikovala kandidátka jako první autorka spolu s dalšími 11 spoluautory v časopise Theoretical and Applied Genetics.

Celkově jsou předkládané teze vynikající prací, která prokazuje jak schopnosti kandidátky, tak vysokou úroveň domovské laboratoře. Úlohu oponenta usnadňuje také skutečnost, že obě prvoautorské publikace již úspěšně prošly náročným recenzním řízením.

Mám jen několik otázek a komentářů.

1. Poměrně rozsáhlý Úvod podává zdařilý a plastický obraz historie, úskalí a nových možností sekvenování velkých rostlinných genomů. Postrádám jen drobný doplněk týkající se možné kombinace PacBio a obdobných metodik generujících dlouhé reads s vysokou chybovostí a Illumina metod produkujících krátké, avšak přesné reads. Existuje řada snah, jak tyto dva postupy spojit, některé vyjádřené i speciálním software. Jistě by stálo za úvahu publikovat Úvod jako samostatné review v anglickém, ale i českém jazyce.

2. Vysoce repetitivní sekvence z genomické mapy č. 350 byla odhalena BioNano mapováním. Existuje nějaký klíč, např. hodnota pokrytí (coverage) malými reads, který by tuto repetici zachytil jen na základě sekvenování (NGS) ?
3. V průběhu práce jistě vznikly i dosud nepublikované výsledky a zajímavé metodické modifikace. Publikace jsou stručné, zejména v metodické oblasti. Může kandidátka doplnit nějaký zajímavý postup, drobné vylepšení, které objevila a které usnadnilo analýzy či pokusy?
4. S předchozím bodem souvisí i tento poslední komentář. Obě publikace jsou prvoautorské, není tedy pochyb o podstatném přínosu kandidátky. Nicméně obě práce mají více než deset spoluautorů. V jaké oblasti kandidátka nejvíce působí? Je její doménou některý laboratorní postup či spíše bioinformatika? V čem ona sama spatřuje svůj největší vklad pro mapování 7DS chromozomu?

V Praze dne 28. Listopadu 2015

RNDr. Helena Štorchová CSc

