

Jarosław Berent

DNASStat wersja 1.2 – program do obsługi bazy danych profili genetycznych oraz do obliczeń biostatystycznych

DNASStat, version 1.2 – a software package for processing genetic profile databases and biostatistical calculations

Z Katedry i Zakładu Medycyny Sądowej Uniwersytetu Medycznego w Łodzi
Kierownik: prof. nadzw. dr hab. n. med. J. Berent

W pracy przedstawiono nową wersję programu DNAStat do obsługi bazy danych profili genetycznych oraz do obliczeń biostatystycznych. Wersja ta – oznaczona numerem 1.2 zawiera – poza wszystkimi możliwościami zawartymi wcześniej w wersji 1.0 także opcję eksportu wyników obliczeń w formacie plików *.xls odczytywanych przez program Microsoft® Office Excel oraz opcję eksportu i importu bazy populacyjnej układów w formacie pliku tekstowego *.txt możliwego do obróbki przez takie programy, jak np. Microsoft® Notatnik lub EditPad.

This paper presents the new DNASStat version 1.2 for processing genetic profile databases and biostatistical calculations. This new version contains, besides all the options of its predecessor 1.0, a calculation-results file export option in .xls format for Microsoft® Office Excel, as well as the option of importing/exporting the population base of systems as .txt files for processing in Microsoft® Notepad or EditPad

WPROWADZENIE

Program DNASStat 1.0 został opracowany w roku 2005 i rozprowadzony nieodpłatnie wśród wszystkich zainteresowanych [1]. Jego opis przedstawiono w poprzedniej publikacji wydrukowanej na łamach Archiwum Medycyny Sądowej i Kryminologii w roku 2006 [2]. Program ten wzbudził duże zainteresowanie i zyskał liczne pochlebne oceny użytkowników. W kursie szkoleniowym, który miał miejsce w dniu 11 stycznia 2006 roku wzięło udział około 70 uczestników [3]. Podczas dyskusji z użytkownikami poja-

wiły się sugestie rozszerzenia możliwości programu przez dodatkowe opcje. Przedstawiona w niniejszej pracy wersja programu 1.2 usuwa kilka drobnych, choć uciążliwych niedogodności, jakie występowały w poprzedniej wersji.

EKSPORT WYNIKÓW OBLICZEŃ

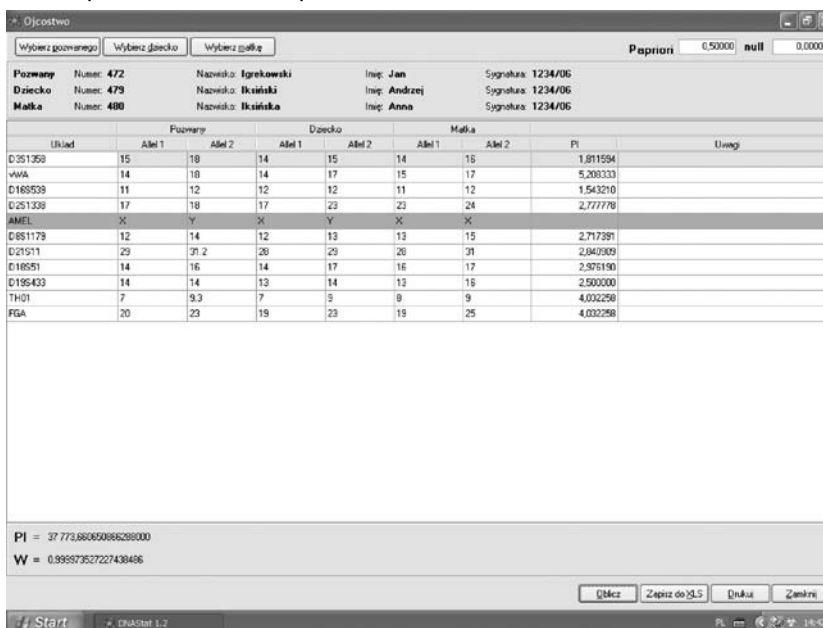
W wersji 1.0 eksport wyników obliczeń był możliwy jedynie w formie wydruku w jednej ustalonej programowo formie, która nie mogła być modyfikowana. Forma ta w praktyce okazała się trudna do bezpośredniego wykorzystania w opiniach. Stąd też wyniki były przepisywane ręcznie, co z kolei mogło być potencjalnym źródłem błędów. Dlatego też dodano do programu opcję eksportu wyników obliczeń w formie plików *.xls odczytywanych przez program Microsoft® Office Excel (ryc. 1 i 2). Taka opcja umożliwia łatwe przeniesienie wyników dokonanych obliczeń do dowolnego edytora tekstów stosowanego w poszczególnych laboratoriach przy pisaniu opinii. Postępowanie takie zmniejsza możliwość popełnienia błędu przez wyeliminowanie ręcznego przepisywania wyników. Warty przypomnienia jest również fakt, że program już w poprzedniej wersji umożliwiał automatyczne importowanie wyników genotypowania przez bezpośrednie wczytanie pliku CODIS tworzego przez sekwenator. Zatem przy dodanej obecnie opcji użytkownik na żadnym etapie pracy nie musi już przepisywać ręcznie genotypów, przy której to czynności potencjalna możliwość zaistnienia błędu jest duża. Program z tą opcją był

oznaczony numerem 1.1 i przedstawiono go zainteresowanym podczas Ogólnopolskiej Konferencji Naukowej: „Opiniowanie sądowo-lekarskie w wybranych działach medycyny sądowej i toksykologii”,

która miała miejsce w Wiśle-Jaworniku w dniach 19-20 października 2006 roku. Obecnie opisywana wersja 1.2 oczywiście również zawiera tę opcję.

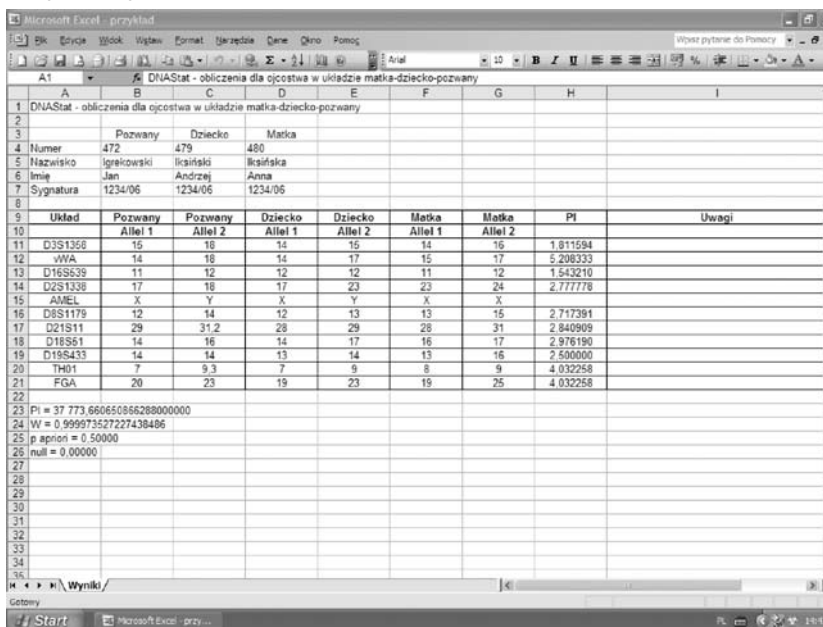
Ryc. 1. Okno programu, z którego po obliczeniach można wyeksportować plik *.xls.

Fig. 1. Program window for post-calculation export of an .xls file.



Ryc. 2. Przykładowy plik *.xls wyeksportowany po obliczeniach.

Fig. 2. Example of an exported post-calculation .xls file..



EKSPORT I IMPORT BAZY POPULACYJNEJ UKŁADÓW

W wersji 1.0 wprowadzenie informacji populacyjnych o układach, takich jak: nazwa układu, częstość

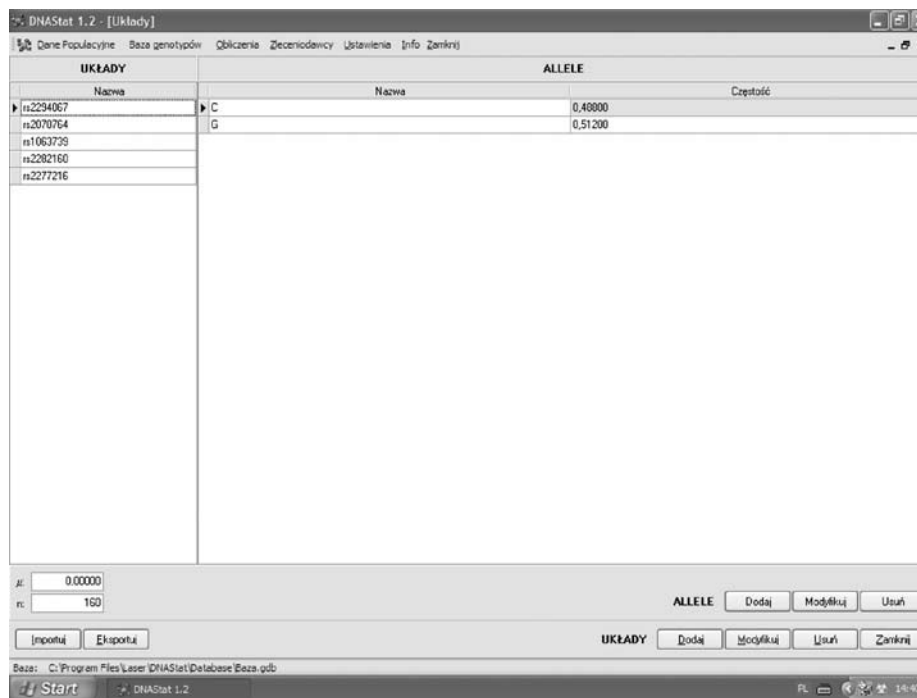
mutacji, wielkość bazy, nazwy poszczególnych alleli i ich częstości było możliwe jedynie z klawiatury poprzez ręczne wpisanie tych danych. Była to procedura żmudna i bardzo utrudniała np. porównanie wyników obliczeń dla różnych baz populacyjnych.

Dlatego też w wersji 1.2 dodano do programu opcję eksportu i importu tych danych w formie pliku tekstowego *.txt możliwego do obróbki przez takie programy, jak np. Microsoft® Notatnik lub EditPad (ryc. 3). Plik taki ma postać: w pierwszej linii gwiazdka, nazwa układu łamane przez częstość mutacji łamane przez wielkość bazy i w kolejnych liniach nazwa allele łamana przez jego częstość. Po liniach zawierających dane o pierwszym układzie następuje jedna linia wolna i w następnych liniach podane są

dane o kolejnych układach. Zaimportowanie pliku z danymi usuwa informacje o układach wcześniej wprowadzone do programu, nie naruszając bazy populacyjnej genotypów (osób). Taka opcja umożliwia szybkie i łatwe modyfikowanie posiadanej bazy np. o nowe allele lub nowe układy oraz prowadzenie obliczeń dla różnych baz. Poniżej (ryc. 4) przedstawiono strukturę przykładowego pliku tekstowego zawierającą dane populacyjne pięciu układów SNP [4].

Ryc. 3. Okno programu, z którego można wyeksportować lub zaimportować plik *.txt.

Fig. 3. Program window for importing/exporting of .txt files.



Ryc. 4. Struktura przykładowego pliku danych.

Fig. 4. Structure of a sample data file.

```
*rs2294067/0,00000/160
C/0,48800
G/0,51200

*rs2070764/0,00000/160
T/0,62500
A/0,37500

*rs1063739/0,00000/160
A/0,48800
C/0,51200

*rs2282160/0,00000/160
G/0,51300
A/0,48700

*rs2277216/0,00000/160
C/0,79400
T/0,20600
```

INSTALACJA AKTUALIZACJI

Plik aktualizacyjny programu „DNASTat 1.2.zip” można nieodpłatnie pobrać ze strony internetowej [5] i po rozpakowaniu wyodrębnić z niego plik „DNASTat 1.2.exe”. Następnie należy wyodrębniony plik skopiować do katalogu C:\Program Files\Laser\DNASTat. Ostatnią czynnością jest zmiana nazwy ikony na pulpicie z „DNASTat 1.0” na „DNASTat 1.2” i odpowiednia zmiana obiektu docelowego we właściwościach tej ikony. Dla prawidłowego działania programu konieczne jest wcześniejsze zainstalowanie wersji 1.0 z dysku instalacyjnego.

PIŚMIENNICTWO

1. Berent J.: DNASTat wersja 1.0 – program do obsługi bazy danych profili genetycznych oraz do

obliczeń biostatystycznych. Program komputerowy. Uniwersytet Medyczny w Łodzi, Łódź 2005.

2. Berent J.: DNASat wersja 1.0 – program do obsługi bazy danych profili genetycznych oraz do obliczeń biostatystycznych. Arch. Med. Sądowej i Kryminol. 2006, 56(1), 15-18.

3. Kurs 05-714-00-201-2006 „Obliczenia biostatystyczne w oparciu o program DNASat 1.0”, Łódź 11.01.2006.

4. Bąbol-Pokora K., Prośniak A., Jacewicz R., Berent J.: Pentapleks SNP – rozkład częstości alleli w populacji centralnej Polski. Arch. Med. Sądowej i Kryminol. – praca w druku.

5. http://wpia.uni.lodz.pl/index.php?go=katedry/katedra_info.php?kat=5170000

Adres do korespondencji:

Prof. nadzw. dr hab. n. med. Jarosław Berent
Kierownik Katedry i Zakładu Medycyny Sądowej
Uniwersytetu Medycznego w Łodzi
ul. Sędziowska 18a
91-304 Łódź
J.Berent@eranet.pl