

Antal Péter – Arany Ádám – Bolgár Bence – Gézsi András – Hajós Gergely  
– Hullám Gábor – Millinghoffer András – Sárközy Péter

## BIOINFORMATIKA LABORATÓRIUM: MOLEKULÁRIS MÉRÉSTECHNIKÁTÓL AZ ORVOSI DÖNTÉSTÁMOGATÁSIG

A bioinformatikai laboratórium jegyzetei (laborsillabuszai) a gyorsan fejlődő bioinformatika új kihívásai mentén haladnak, támaszkodva a Bioinformatika jegyzet megfelelő fejezeteiben található elméleti összefoglalókra és a virtuálislabor-koncepcióban kidolgozott eszközökre. Az első fejezet a biobankok és laboratóriumi információs rendszerek világát mutatja be. Ezt követően a genetikai méréstechnikák eszközeit foglalja össze két fejezet. A kísérlettervezést, különösen a többlépcsős és nagy mennyiségű, heterogén adat és tudás fúzióján alapuló kísérlettervezést mutatja be a következő fejezet. Az adatok hatékony megosztása és a nagy számításigény miatt egyre fontosabbá válnak az általánosan elérhető, közmű jellegű informatikai szolgáltatások, amelyek működését tekinti át a következő fejezet. A genetikai és génexpressziós adatok standard elemzését egy-egy fejezet ismerteti, amit egy rendszerbiológiai ihletésű többváltozós biomarker-elemzési módszertan bemutatása és egy Bayes-i oksági következtetési metodológia áttekintése követ. A biomarkerek felhasználását döntéstámogató rendszerek építése és adaptív döntéstámogató rendszerek létrehozása kapcsán is bemutatjuk egy-egy fejezetben. Az utolsó két sillabusz egy gyógyszerkutatási megközelítést és egy metagenomika elemzési munkafolyamatot foglal össze.

Kulcsszavak: biobankok, laboratóriumi információs rendszerek, genotipizálás, új generációs szekvenálási módszerek, kísérlettervezés, munkafolyamat-rendszerek, asszociációs elemzések, biomarker-elemzések, orvosi döntéstámogató rendszerek, gyógyszerkutató, metagenomika.

Budapesti Műszaki és Gazdaságtudományi Egyetem és Semmelweis Egyetem



**ÚJ SZÉCHENYI TERV**

Typotex Kiadó  
2014

COPYRIGHT: © 2014–2019, Antal Péter, Arany Ádám, Bolgár Bence, Gézsi András, Hajós Gergely, Hullám Gábor, Millinghoffer András, Sárközy Péter, Budapesti Műszaki és Gazdaságtudományi Egyetem, Semmelweis Egyetem

Creative Commons NonCommercial-NoDerivs 3.0 (CC BY-NC-ND 3.0)

A szerző nevének feltüntetése mellett nem kereskedelmi céllal szabadon másolható, terjeszthető, megjelentethető és előadható, de nem módosítható.

Szakmai lektorok: Molnár Viktor, Antos András

ISBN 978 963 279 182 1

Készült a Typotex Kiadó gondozásában

Felelős vezető: Votisky Zsuzsa

Készült a TÁMOP-4.1.2/A/1-11/1-2011-0079 számú, „Konzorcium a biotechnológia aktív tanulásáért” című projekt keretében.

Nemzeti Fejlesztési Ügynökség  
www.ujszochenytterv.gov.hu  
06 40 638 638



A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

# Tartalomjegyzék

<b>1. Biobankok. Laboratóriumi információs rendszerek</b>	<b>7</b>
1.1. Bevezetés . . . . .	7
1.1.1. Biobankok . . . . .	7
1.1.2. Laboratóriumi információs rendszerek . . . . .	7
1.2. A LIMS-ek képességei . . . . .	8
1.2.1. Alapvető funkciók . . . . .	8
1.2.2. További funkciók . . . . .	9
1.3. LIMS: esettanulmány . . . . .	10
1.4. Ellenőrző kérdések . . . . .	14
<b>2. DNS rekombináns méréstechnológiák, zaj- és hibamodellek</b>	<b>15</b>
<b>3. Virtuális genotipizálási laboratóriumi segédlet</b>	<b>19</b>
<b>4. Kísérlettervezés: az alapoktól a tudásgazdag kiterjesztésekig</b>	<b>27</b>
<b>5. Bioinformatikai munkafolyamat-rendszerek</b>	<b>35</b>
5.1. Adat és modell konstruálása . . . . .	35
5.2. A BMLA-elemzést leíró fájl . . . . .	35
5.3. Futtatást a HTCCondor rendszerben . . . . .	37
5.4. Nyers futási eredmények aggregálása . . . . .	38
5.5. Ellenőrző kérdések . . . . .	40
<b>6. Genetikai asszociációs vizsgálatok standard elemzése laborgyakorlat</b>	<b>41</b>
6.1. Bevezetés . . . . .	41
6.2. Hardy–Weinberg-egyenlőség vizsgálata . . . . .	41
6.3. Standard asszociációs tesztek . . . . .	42
6.4. Haplotípus-asszociáció vizsgálata . . . . .	44
6.4.1. Kapcsoltság . . . . .	44
6.4.2. Haplotípusblokkok meghatározása . . . . .	46
6.4.3. Asszociációs tesztek . . . . .	46
6.4.4. Permutációs tesztek . . . . .	48

<b>7. Génexpressziós adatok standard asszociációs elemzése</b>	<b>50</b>
7.1. Bevezetés . . . . .	50
7.1.1. Az adat . . . . .	50
7.2. Az elemzéshez szükséges további csomagok . . . . .	51
7.3. Az adatok letöltése . . . . .	52
7.4. Minőségi ellenőrzések . . . . .	53
7.5. Az adathalmaz szűrése . . . . .	54
7.6. Differenciálisan expresszáló gének meghatározása . . . . .	55
<b>8. Bayes-i, rendszerszintű biomarker-elemzés</b>	<b>59</b>
8.1. Bevezető . . . . .	59
8.2. Emlékeztető kérdések . . . . .	61
8.3. Feladatok . . . . .	61
8.4. MBS poszteriorok utófeldolgozása és megjelenítése . . . . .	62
8.4.1. Feltételes MBS poszteriorok megjelenítése a modell struktúrára vetítve	62
8.4.2. MBS és k-MBS poszteriorok megjelenítése részhalmazhálón . . . . .	62
8.4.3. A relevanciafa . . . . .	63
8.4.4. A relevancia-interakció . . . . .	63
<b>9. Heterogén biológiai adatok fúziós elemzése és értelmezése</b>	<b>68</b>
<b>10. Bayes-i oksági elemzés</b>	<b>79</b>
10.1. Bevezetés . . . . .	79
10.2. Emlékeztető kérdések . . . . .	81
10.3. Feladatok . . . . .	81
10.4. Feltételes MBG poszteriorok megjelenítése a modellstruktúrára vetítve . . . . .	81
10.5. Páronkénti poszteriorok megjelenítése a modellstruktúrára vetítve . . . . .	81
<b>11. Tudásmérnöki technikák alkalmazása döntési hálónál</b>	<b>84</b>
11.1. Bevezetés . . . . .	84
11.2. Kérdések/Emlékeztetők . . . . .	84
11.3. Tudásmérnöki technikák Bayes-hálókhoz . . . . .	85
11.4. Feladatok . . . . .	85
11.5. Bayes-háló-modellek szerkesztése . . . . .	85
11.5.1. Új modell létrehozása . . . . .	85
11.5.2. Létező modell megnyitása . . . . .	86
11.5.3. Modellstruktúra szerkesztése . . . . .	86
11.5.4. Csomópontok típusainak szerkesztése . . . . .	86
11.5.5. Csomópontok csoportjainak szerkesztése . . . . .	86
11.5.6. Csomópontok szerkesztése . . . . .	87
11.5.7. Annotáció és magyarázatkeresés . . . . .	91
11.6. Következtetés Bayes-hálókban . . . . .	92
11.6.1. Evidenciák/döntések bevitele . . . . .	93

11.6.2. Marginális eloszlások számítása . . . . .	94
11.6.3. Következtetés érzékenységeinek vizsgálata . . . . .	95
11.7. Strukturális információk megjelenítése . . . . .	96
11.7.1. Élek megjelenítése . . . . .	97
11.7.2. Húrgráf megjelenítése . . . . .	98
11.7.3. Klikkfa megjelenítése . . . . .	98
<b>12. Döntéstámogató rendszerek tanulása</b>	<b>101</b>
12.1. Bevezető . . . . .	101
12.2. Kérdések/Emlékeztetők . . . . .	101
12.3. Feladatok . . . . .	101
12.4. Becslések torzítása . . . . .	102
12.5. Paraméterek tanulása . . . . .	103
12.5.1. Adatfájlok formátuma . . . . .	103
12.5.2. Paraméterek beállítása adatfájl alapján . . . . .	103
12.6. Modellkiértékelés adatfájl alapján . . . . .	103
12.7. Struktúratanulás . . . . .	104
<b>13. Virtuális szűrés kernel-módszerekkel</b>	<b>108</b>
13.1. Bevezetés . . . . .	108
13.2. A referencia-vegyületelemkészítése . . . . .	108
13.3. A kernelek elkészítése . . . . .	110
13.4. Egy-osztályos prioritizálás . . . . .	111
13.5. Kvantitatív struktúra-hatás összefüggés . . . . .	111
13.6. Kérdések . . . . .	113
<b>14. Metagenomika</b>	<b>115</b>
14.1. Bevezetés . . . . .	115
14.2. Előfeldolgozás . . . . .	115
14.3. Adatelemzés . . . . .	119
14.3.1. Taxonómiai alapegységek meghatározása . . . . .	119
14.3.2. Alfa-diverzitás . . . . .	120
14.3.3. Béta-diverzitás . . . . .	123
14.4. Ellenőrző kérdések . . . . .	125