

6. Literaturverzeichnis

- Becker KG**, White SL, Muller J, Engel J (2000) BBID: the biological biochemical image database. *Bioinformatics* **16**(8):745-6.
- Benson DA**, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2004) *Nucleic Acids Res.* **32**:D23-6.
- Benton, D** (1996) Bioinformatics – principles and potential of a new multidisciplinary tool. *TibTech* **14**:261-72.
- Berrar D**, Dubitzky W, Solinas-Toldo S, Bulashevskaya S, Granzow M, Conrad C, Kalla J, Lichter P, Eils R (2001) A database system for comparative genomic hybridization analysis. *IEEE Eng Med Biol Mag.* **20**(4):75-83.
- Blake JA**, Richardson JE, Bult CJ, Kadin JA, Eppig JT, and the members of the Mouse Genome Database Group (2003) MGD: The Mouse Genome Database. *Nucleic Acids Res* **31**: 193-195.
- Bobo, M** (2002) Improving Functional Analysis of Microarray Gene Expression Data. Masters-Arbeit Oregon Health and Science University, USA.
- Brazma A**, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N, Holloway E, Kapushesky M, Kemmeren P, Lara GG, Oezcimen A, Rocca-Serra P, Sansone S (2003) ArrayExpress - a public repository for microarray gene expression data at the EBI. *Nucleic Acids Research* **31**:68-71.
- Castillo-Davis CI**, Hartl DL (2003) GeneMerge - post-genomic analysis, data mining, and hypothesis testing. *Bioinformatics* **19**(7):891-2.
- Chen, PP** (1976) Entity-Relationship Model: Towards a Unified View of Data, *ACM Transactions on Database Systems* **1**, 9-36.
- Codd, EF** (1970) A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, **13**(6):377-387.
- Curwen V**, Eyraas E, Andrews TD, Clarke L, Mongin E, Searle SMJ, Clamp M (2004) The Ensembl Automatic Gene Annotation System *Genome Res* **14**:942-50.

- Dooley TP**, Reddy SP, Wilborn TW, Davis RL (2003) Biomarkers of human cutaneous squamous cell carcinoma from tissues and cell lines identified by DNA microarrays and qRT-PCR. *Biochem. Biophys. Res. Commun.* **306**:1026-1036.
- du Manoir S**, Speicher MR, Joos S, Schrock E, Popp S, Dohner H, Kovacs G, Robert-Nicoud M, Lichter P, Cremer T (1993) Detection of complete and partial chromosome gains and losses by comparative genomic in situ hybridization. *Hum Genet.* **90**(6):590-610.
- Eckert RL**, Broome AM, Ruse M, Robinson N, Ryan D, Lee K. (2004) S100 proteins in the epidermis. *J Invest Dermatol.* **123**(1):23-33.
- Edgar R**, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, **30**(1):207-10.
- Ermolaeva O**, Rastogi M, Pruitt KD, Schuler GD, Bittner ML, Chen Y, Simon R, Meltzer P, Trent JM & Boguski MS (1998) Data management and analysis for gene expression arrays. *Nature Genetics* **20**:19-23.
- Etzold T**, Ulyanow A and Argos P (1996). SRS: Information Retrieval System for Molecular Biology Data Banks. *Methods Enzymol.* **266**:114-128.
- eWEEK:** Vergleichstest führender Datenbank-Managementsysteme.
<http://www.eweek.com/article2/0,4149,293,00.asp>
- The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team** (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420**:563-573,
- Fellenberg K**, Hauser NC, Brors B, Hoheisel JD, Vingron M (2002) Microarray data warehouse allowing for inclusion of experiment annotations in statistical analysis. *Bioinformatics* **18**(3):423-33.
- Fiegler H**, Carr P, Douglas EJ, Burford DC, Hunt S, Scott CE, Smith J, Vetrie D, Gorman P, Tomlinson IP, Carter NP (2003) DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer* **36**(4):361-74.
- Fritz B**, Schubert F, Wrobel G, Schwaenen C, Wessendorf S, Nessling M, Korz C, Rieker RJ, Montgomery K, Kucherlapati R, Mechttersheimer G, Eils R, Joos S,

- Lichter P (2002) Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. *Cancer Res.* **62**(11):2993-8.
- GeneOntology Consortium** (2001) Creating the gene ontology resource: design and implementation. *Genome Res*, **11**:1425-33.
- Gilbert DG** (2002) euGenes: a eukaryote genome information system. *Nucleic Acids Res*, **30**:145-148.
- Gollub J**, Ball CA, Binkley G, Demeter J, Finkelstein DB, Hebert JM, Hernandez-Boussard T, Jin H, Kaloper M, Matese JC, Schroeder M, Brown PO, Botstein D, Sherlock G (2003) The Stanford Microarray Database: data access and quality assessment tools. *Nucleic Acids Res.* **31**(1):94-6.
- Golub, TR**, D. K. Slonim,DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, BloomPeld CD, Lander ES (1999) Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. *Science*, **286**:531-537.
- Haas LM**, Schwarz PM, Kodali P, Kotlar E, Rice J and Swope WC (2001) DiscoveryLink: A System for Integrated Access to Life Sciences Data Sources. *IBM Systems Journal* **40**(2): 489-511.
- Haas S**, Vingron M, Poustka A, and Wiemann S (1998) Primer design for large scale sequencing. *Nucleic Acids Res* **26**(12), 3006.3012.
- Herrero J**, Vaquerizas JM, Al-Shahrour F, Conde L, Mateos A, Diaz-Uriarte JS, Dopazo J. (2004) New challenges in gene expression data analysis and the extended GEPAS. *Nucleic Acids Res.* **32**(Web Server Issue):W485-W491.
- Holloway AJ**, van Laar RK, Tothill RW, Bowtell DDL (2002) Options available – from start to finish – for obtaining data from DNA microarrays II. *Nat Gen* **32** (Suppl): 481-489.
- Hosack DA**, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol*, **4**(10):R70.
- Hubbard T**, Barker D, Birney E, Cameron G, Chen Y, Clark L, Cox T, Cuff J, Curwen V, Down T, Durbin R, Eyraas E, Gilbert J, Hammond M, Huminiecki L, Kasprzyk A, Lehvaslaiho H, Lijnzaad P, Melsopp C, Mongin E, Pettett R, Pockock M, Potter S, Rust A, Schmidt E, Searle S, Slater G, Smith J, Spooner W, Stabenau A, Stalker J,

- Stupka E, Ureta-Vidal A, Vastrik I, Clamp M (2002) The Ensembl genome database project. *Nucleic Acids Res*, **30**(1):38-41.
- Hummerich L**, Müller R, Hess J, Kokocinski F, Hahn M, Fürstenberger G, Mauch C, Lichter P and Angel P (Eingereicht) Multistage Skin Cancer Development Involves Members from the Epidermal Differentiation Complex. Eingereicht.
- Ishkanian AS**, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, Ling V, MacAulay C, Lam WL (2004) A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet*. **36**(3):299-303.
- Ito N**, Hasegawa R, Imaida K, Hirose M, Asamoto M, Shirai T (1995) Concepts in multistage carcinogenesis. *Crit. Rev. Oncol. Hematol.*, **21**: 105-133.
- Joos S**, Scherthan H, Speicher MR, Schlegel J, Cremer T, Lichter P (1993) Detection of amplified DNA sequences by reverse chromosome painting using genomic tumor DNA as probe. *Hum Genet*. **90**(6):584-9.
- Jurka J** (2000) Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet* **16**(9):418–420.
- Kallioniemi A**, Kallioniemi O-P, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D (1992) Comparative Genomic Hybridization for Molecular Cytogenetic Analysis of Solid Tumors. *Science* **258**: 818-821.
- Karolchik D**, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, Roskin KM, Schwartz M, Sugnet CW, Thomas DJ, Weber RJ, Haussler D, Kent WJ (2003) The UCSC Genome Browser Database. *Nucleic Acids Res*. **31**(1):51-4.
- Kasprzyk A**, Keefe D, Smedley D, London D, Spooner W, Melsopp C, Hammond M, Rocca-Serra P, Cox T, Birney E (2004) EnsMart: a generic system for fast and flexible access to biological data. *Genome Res* **14**:160-9.
- Khatri P**, Draghici S, Ostermeier GC, Krawetz SA (2002) Profiling gene expression using Onto-express. *Genomics* **79**(2):266-70.
- Kokocinski F**, Wrobel G, Hahn M, Lichter P (2003) QuickLims: Facilitating the data management for DNA-microarray production. *Bioinformatics* **19** (2):283-284.

- Korshunov A**, Neben K, Wrobel G, Tews B, Benner A, Hahn M, Golanov A, Lichter P (2003) Gene expression patterns in ependymomas correlate with tumor location, grade, and patient age. *Am J Pathol.* **163**(5):1721-7.
- Korz C**, Pscherer A, Benner A, Mertens D, Schaffner C, Leupolt E, Dohner H, Stilgenbauer S, Lichter P (2002) Evidence for distinct pathomechanisms in B-cell chronic lymphocytic leukemia and mantle cell lymphoma by quantitative expression analysis of cell cycle and apoptosis-associated genes. *Blood.* **99**(12):4554-61.
- Lennon G**, Auffray C, Polymeropoulos M, Soares MB (1996) The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics*, **33**(1):151-2.
- Leser U, Rieger P** (2003) Integration molekularbiologischer Daten.
- Lipshutz RJ**, Morris D, Chee M, Hubbell E, Kozal MJ, Shah N, Shen N, Yang R, Fodor SP (1995) Using Oligonucleotide Probe Arrays to Access Genetic Diversity. *Biotechniques* **19**(3):442–447.
- Mangalam H**, Stewart J, Zhou J, Schlauch K, Waugh M, Chen G, Farmer AD, Colello G, Weller JW (2001) GeneX: An Open Source gene expression database and integrated tool set. *IBM Systems J*, **40**(2).
- Marshall OJ** (2004) PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* [Epub ahead of print].
- Neben K**, Giesecke C, Schweizer S, Ho AD, Kramer A (2003 a) Centrosome aberrations in acute myeloid leukemia are correlated with cytogenetic risk profile. *Blood.* **101**(1):289-91.
- Neben K**, Tews B, Wrobel G, Hahn M, Kokocinski F, Giesecke C, Krause U, Ho AD, Krämer A, Lichter P (2003 b) Gene Expression Pattern in Acute Myeloid Leukemia Correlate with Centrosome Changes. *Oncogene* **23** (13), 2379-84.
- Neben K**, Korshunov A, Benner A, Wrobel G, Hahn M, Kokocinski F, Golanov A, Joos S, Lichter P (2004) Microarray-Based Screening for Molecular Markers in Medulloblastoma Revealed STK15 as Independent Predictor for Survival. *Cancer Res.* **64**(9):3103-11.
- Oberlé V** (2000) Data Mining: eine Einführung. <http://vincent.oberle.com/data-mining.pdf>.

- Online Mendelian Inheritance in Man, OMIM (TM)** (2000) McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD), <http://www.ncbi.nlm.nih.gov/omim>.
- Pattyn F**, Speleman F, De Paepe A, and Vandesompele J (2003) RTPrimerDB: the real-time PCR primer and probe database. *Nucleic Acids Res* **31**(1), 122.123.
- Pease AC**, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP (1994) Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* **91**(11):5022-6.
- Perou CM**, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours *Nature* **406**(6797):747-52.
- Persidis A** (1999) Bioinformatics. *Nat Biotechnol* **17**:828-30.
- Pinkel D**, Seagraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung B, Gray JW, Albertson DG (1998) High resolution analysis of DNA copy-number variations using comparative genomic hybridization to microarrays. *Nature Genetics* **20**: 207-211..
- PRNewswire**: Vergleich führender Webserver. <http://www.prnewswire.com/cgi-bin/stories.pl?ACCT=SVBIZINK3.story&STORY=/www/story/05-11-2004/0002172126>.
- Pruitt KD**, Katz KS, Sicotte H, Maglott DR (2000) Introducing RefSeq and LocusLink: curated human genome resources at the NCBI. *Trends Genet*, **16**(1):44-7.
- Rebhan M**, Chalifa-Caspi V, Prilusky J, Lancet D (1997) GeneCards: encyclopedia for genes, proteins and diseases. Weizmann Institute of Science, Bioinformatics Unit and Genome Center (Rehovot, Israel), <http://bioinformatics.weizmann.ac.il/cards>.
- Rosenwald A**, Alizadeh AA, Widhopf G, Simon R, Davis RE, Yu X, Yang L, Pickeral OK, Rassenti LZ, Powell J, Botstein D, Byrd JC, Grever MR, Cheson BD, Chiorazzi N, Wilson WH, Kipps TJ, Brown PO, Staudt LM (2001) Relation of Gene Expression Phenotype to Immunoglobulin Mutation Genotype in B Cell Chronic Lymphocytic Leukemia. *J Ex Med* **194**(11):1639–1647.

- Rozen S and Skaletsky H** (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**:365-386.
- Schlingemann J**, Hess J, Wrobel G, Breitenbach U, Gebhardt C, Steinlein P, Kramer H, Furstenberger G, Hahn M, Angel P, Lichter P (2003) Profile of gene expression induced by the tumour promotor TPA in murine epithelial cells. *Int. J. Cancer* **104**:699-708.
- Schwaenen C**, Nessling M, Wessendorf S, Salvi T, Wrobel G, Radlwimmer B, Kestler HA, Haslinger C, Stilgenbauer S, Dohner H, Bentz M, Lichter P (2004) Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. *Proc Natl Acad Sci U S A*. **101**(4):1039-44.
- Shena M**, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**:467-470.
- Solinas-Toldo S**, Lampel S, Benner A, Cremer T, Lichter P (1997) Matrix-based comparative genomic hybridization: Biochips to screen for genomic imbalances. *Genes Chromosomes & Cancer* **20**: 399-407.
- Southern EM** (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**: 503-518.
- Stabenau A**, McVicker G, Melsopp C, Proctor G, Clamp M, Birney E (2004) The Ensembl Core Software Libraries *Gen Res* **14**:929-33.
- Stajich J E**, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, Fuellen G, Gilbert JGR, Korf I, Lapp H, Lehvaslaiho H, Matsalla C, Mungall CJ, Osborne BI, Pocock MR, Schattner P, Senger M, Stein LD, Stupka E, Wilkinson MD, and Birney E (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* **12**(10), 1611-1618.
- Stratowa C**, Loffler G, Lichter P, Stilgenbauer S, Haberl P, Schweifer N, Dohner H, Wilgenbus KK (2001) CDNA microarray gene expression analysis of B-cell chronic lymphocytic leukemia proposes potential new prognostic markers involved in lymphocyte trafficking. *Int J Cancer* **91**(4):474-80.
- Strausberg RL**, Buetow KH, Emmert-Buck M, Klausner R (2000) The Cancer Genome Anatomy Project: building an annotated gene index. *Trends in Genetics*, **16**, 103-106.

- Thareau V**, Déhais P, Serizet C, Hilson P, Rouzé P, Aubourg S (2003) Automatic design of gene-specific sequence tags for genome-wide functional studies. *Bioinformatics* **19**(17):2191-98.
- Velculescu VE**, Zhang L, Vogelstein B, Kinzler KW (1995) Serial Analysis of Gene Expression. *Science* **270**(5235):484–487.
- von Mering C**, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. (2003) STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* **31**(1):258-61.
- Wang X**, Seed B (2003) A PCR primer bank for quantitative gene expression analysis. *Nucleic Acids Res* **31**(24).
- Wessendorf S**, Fritz B, Wrobel G, Nessling M, Lampel S, Goettel D, Kuepper M, Joos S, Hopman T, Kokocinski F, Dohner H, Bentz M, Schwaenen C, Lichter P (2002) Automated screening for genomic imbalances using matrix-based comparative genomic hybridization. *Lab Invest.* **82**(1):47-60.
- Wessendorf S**, Schwaenen C, Kohlhammer H, Kienle D, Wrobel G, Barth TF, Nessling M, Moller P, Dohner H, Lichter P, Bentz M (2003) Hidden gene amplifications in aggressive B-cell non-Hodgkin lymphomas detected by microarray-based comparative genomic hybridization. *Oncogene.* **22**(9):1425-9.
- Wilgenbus KK**, Lichter P (1999) DNA chip technology ante portas. *J Mol Med* **77**:761-768.
- Wittwer CT**, Fillmore GC, Hillyard DR (1989) Automated polymerase chain reaction in capillary tubes with hot air. *Nucleic Acids Res.* **17**(11):4353-7.
- Wrobel G**, Schlingemann J, Hummerich L, Kramer H, Lichter P, and Hahn M (2003) Optimization of high-density cDNA-microarray protocols by 'design of experiments'. *Nucleic. Acids. Res.* **31**:e67.
- Wrobel G** (2004) Aufbau, Validierung und Verwendung eines cDNA-Microarray-Systems für die Expressionsanalyse humaner Meningiome. Dissertation, Universität Heidelberg, Mathematisch-Naturwissenschaftliche Gesamtfakultät.
- Wrobel G**, Kokocinski F, Lichter P (2004) AutoPrime: selecting primers for expressed sequences. *Genome Biology* **5**:P11 (<http://genomebiology.com/2004/5/5/P11>).

Wrobel G, Roerig P, Kokocinski F, Neben K, Hahn M, Reifenberger G, Lichter P (eingereicht) Gene Expression profiling of Benign, Atypical and Anaplastic Meningiomas Identifies Novel Genes Associated with Meningioma Progression. Eingereicht.

Zeeberg BR, Feng W, Wang G, Wang MD, Fojo AT, Sunshine M, Narasimhan S, Kane DW, Reinhold WC, Lababidi S, et al. (2003) GoMiner: A Resource for Biological Interpretation of Genomic and Proteomic Data. *Genome Biol* **4**(4):R28.

Zhu H, Snyder M (2003) Protein chip technology. *Curr Opin Chem Biol*. **7**(1):55-63.

7. Anhang

7.1. Erläuterung von Abkürzungen und Fachtermini

Annotation: Zusatzinformationen zu einem bestimmten Objekt, z.B. genomische Lokalisierungs-Information zu einem Gennamen

API: Definierte Schnittstelle eines Systems, die von anderen Programmen zur Anbindung genutzt werden kann (*Application Programmer Interface*)

ASCII: Standard Zeichensatz (des romanischen Alphabets) (*American Standard Code for Information Interchange*)

B, MB, GB: Computer-Speichereinheiten (Byte, Megabyte, Gigabyte)

b, kb, Mb: Nukleinsäuren-Kettenlänge (Basenpaare, Kilo-Basenpaare, Mega-Basenpaare)

Client: Computer, der die von einem Server zur Verfügung gestellten Dienste oder Daten in einem Netzwerk nutzt.

CPAN: Öffentliches Archiv von Perl-Modulen (<http://cpan.org>; *Comprehensive Perl Archive Network*)

Datenbank / DB: Computersystem, welches zur effizienten Verwaltung große Datenmengen in strukturierter Form speichert und mit speziellen Mechanismen zur Abfrage bereithält.

DB: Datenbank

DNA, RNA, cDNA, mRNA: Nukleinsäure-basierte Moleküle der Erbsubstanz an einer Zucker-Phosphat-Kette

DKFZ: Deutsches Krebsforschungszentrum Heidelberg

FTP: Definition eines Protokolls zur elektronischen Datei-Übermittlung (*File Transfer Protocol*)

GO: GeneOntology; hierarchisches System (in Form eines azyklischen ungerichteten Graphen) von Annotationsbegriffen der Kategorien Molekulare Funktion, Biologischer Prozess und Zelluläre Lokalisierung.

JPG: Dateiformat zur komprimierten Speicherung von Graphikdaten (*Joint Expert Group-Format*)

LIMS: Computersystem zur Verwaltung von Labordaten und zur Kontrolle von Prozessabläufen (*Laboratory Information and Management System*)

MHz: Taktfrequenz, mit der Computerkomponenten arbeiten und Daten weiterleiten können (Megahertz)

PAC: Künstlicher Vektor, in den 100-200 kb DNA kloniert werden können (*Plasmid Artificial Chromosome*)

PCR: Polymerase-Kettenreaktion zur Vervielfältigung von Nukleinsäuren (*Polymerase Chain Reaction*)

RAID: Verwaltungs- und Zugriffssystem von mehreren Festplattenspeichern an einem Computer (*Redundant Array of Independent Disks*)

RAM: Beschreibbarer Arbeitsspeicher eines Computers (*Random-Access-Memory*)

RZPD: Ressourcenzentrum des dt. Humanen Genomprojekts in Berlin und Heidelberg (*Ressource Center / Primary Database*)

Server: Computer, der bestimmte Dienste, bzw. Daten in einem Netzwerk (Intranet oder Internet) zur Verfügung stellt.

XML: Flexible Definitionssprache zum Austausch von Daten zwischen heterogenen Anwendungen und Geräten (*eXtensible Markup Language*).

7.2. Eigene Publikationen

7.2.1. Publikationen in Fachzeitschriften

- Kokocinski F, et al.** (2004) *Functional interpretation of high-throughput experiments with FACT*. In Vorbereitung.
- Hummerich L, Müller R, Hess J, **Kokocinski F**, Hahn M, Fürstenberger G, Mauch C, Lichter P and Angel P (2004) *Multistage Skin Cancer Development Involves Members from the Epidermal Differentiation Complex*. Eingereicht.
- Wrobel G, Roerig P, **Kokocinski F**, Neben K, Hahn M, Reifenberger G, Lichter P (2004) *Microarray-based Gene Expression Profiling of Benign, Atypical and Anaplastic Meningiomas Identifies Novel Genes Associated with Meningioma Progression*. Eingereicht.
- Florin L, Hummerich L, Dittrich B, **Kokocinski F**, Wrobel G, Gack S, Schorpp-Kistner M, Werner S, Hahn M, Lichter P, Szabowski A, Angel P (2004) Identification of novel AP-1 target genes in fibroblasts regulated during cutaneous wound healing. *Oncogene*. Im Druck.
- Neben K, Korshunov A, Benner A, Wrobel G, Hahn M, **Kokocinski F**, Golanov A, Joos S, Lichter P (2004) *Microarray-Based Screening for Molecular Markers in Medulloblastoma Revealed STK15 as Independent Predictor for Survival*. *Cancer Res.* **64**(9):3103-11.
- Wrobel G, **Kokocinski F**, Lichter P (2004) *AutoPrime: Selecting Primers for Expressed Sequences*. *Genome Biology*, **5**:P11 (<http://www.genome-biology.com>).
- Neben K, Tews B, Wrobel G, Hahn M, **Kokocinski F**, Giesecke C, Krause U, Ho AD, Krämer A, Lichter P (2004) *Gene Expression Pattern in Acute Myeloid Leukemia Correlate with Centrosome Changes*. *Oncogene* **23** (13), 2379-84.
- Kokocinski F**, Wrobel G, Hahn M, Lichter P (2003) *QuickLims: Facilitating the data management for DNA-microarray production*. *Bioinformatics* **19** (2), 283-284.
- Fritz B, **Kokocinski F**, Schlingemann J, Hahn M (2003) *Anwendungen der DNA-Chiptechnologie in der Krebsforschung*. *BioSpektrum* **1**, 78-83.
- Wessendorf S, Fritz B, Wrobel G, Nessling M, Lampel S, Goettel D, Kuepper M, Joos S, Hopman T, **Kokocinski F**, Döhner H, Bentz M, Schwaenen C, Lichter P (2002) *Automated screening for genomic imbalances using matrix-based comparative genomic hybridization (matrix-CGH)*. *Lab. Invest.* **82**, 47-60.

7.2.2. Beiträge zu internationalen Konferenzen

Kokocinski F, Wrobel G, Toedt T, Lichter P (2003) *Facilitating the interpretation of high-throughput experiments*. European Conference on Computational Biology 2003, Paris / Frankreich.

Toedt G, Wrobel G, **Kokocinski F**, Lichter P (2003) *ChipYard*. European Conference on Computational Biology 2003, Paris / Frankreich.

Schlingemann J, Hummerich L, Hess J, Wrobel G, Breitenbach U, **Kokocinski F**, Fürstenberger G, Hahn M, Angel P, Lichter P (2003) *Expression profiling of murine epithelial cells*. Statusseminar Chiptechnologien, Frankfurt.

Wrobel G, Büschges R, **Kokocinski F**, Weber R, Neben K, Hahn M, Reifenberger G, Lichter P (2003) *Expression profiling in benign, atypical and anaplastic meningioma*. Statusseminar Chiptechnologien, Frankfurt.

Kokocinski F, Wrobel G, Lichter P (2002) *Process management for the chip fabrication*. European Conference on Computational Biology 2002, Saarbrücken

Hummerich L, Schlingemann J, Hess J, Breitenbach U, Wrobel G, **Kokocinski F**, Fürstenberger G, Hahn M, Angel P, Lichter P (2002) *Expression Profiling of Epithelial Cells in Skin Cancerogenesis*. NGFN, Berlin.

Kokocinski F, Wrobel G, Lichter P (2002) *Establishment of a DNA-chip database network*. Statusseminar Chiptechnologien, Frankfurt.

Zusammenfassung

Der Einsatz von Hochdurchsatz-Methoden wie *matrixCGH* und *Expression Profiling* mit DNA-Microarrays in der molekular-medizinischen Forschung erlaubt die parallele Analyse von biologischen Messwerten in einer bisher unerreichten Größenordnung. Um die anfallende Menge an Daten optimal nutzen zu können bedarf es spezialisierter Software-Lösungen.

Für das Management einer Klonsammlung, welche den Ausgangspunkt der Experimente darstellt, wurde das *CloneBase*-System entwickelt. Es handelt sich um eine Datenbank mit Internetzugangsseiten und einer automatisierten Aktualisierung der Daten. Es führt damit die Archivierung, Annotation und Informationsbereitstellung für sämtliche Klone durch. Von den Nukleinsäurefragmenten selektierter Klone werden mit Hilfe von Robotern Microarrays produziert. Um den komplexen Herstellungsprozess steuern und überwachen zu können, wurde ein Labordatensystem (*Laboratory Information Management System, LIMS*) entwickelt. *QuickLIMS* arbeitet mit den per Barcode identifizierten Mikrotiterplatten, die die Klone enthalten, indem es sie schrittweise durch ein definiertes Protokoll führt. Der Prozessablauf ist flexibel gespeichert, die Benutzeroberfläche wird darauf basierend dynamisch generiert. Mit dem System kann eine größtmögliche Automatisierung und eine Rückverfolgung eventueller Fehler bei der Microarrayherstellung erreicht werden. Für die Interpretation der experimentellen Ergebnisse von dieser Art von Hochdurchsatz-Methoden stellt die Integration unterschiedlicher Informationsquellen eine Schlüsselrolle dar. Das Analysesystem *FACT (Flexible Annotation and Correlation Tool)* wurde für die explorative Analyse der Ergebnisse mit Hilfe von Annotationsdaten und die Zusammenführung von relevanten Informationen aus heterogenen Datenquellen entwickelt. Daten aus unterschiedlichen experimentellen Systemen, z.B. genomische Veränderungen aus *matrixCGH*-Versuchen und Expressionswerte von mit cDNA-Microarrays untersuchten Genen, können mit Hilfe von spezialisierten Einzelmodulen in Zusammenhang gebracht werden. Dies wird durch die flexible Speicherung der Datendefinitionen und die Transformation auf eine gemeinsame Datenbasis erreicht. In gleicher Art ist es möglich, unterschiedliche Annotationsquellen und neue Analysemethoden in das System einzubinden. Mit der Anwendung *AutoPrime* wird schließlich auch die Verifikation der erzielten Microarray-Ergebnisse mit der Methode der RQ-PCR zu einem weiteren Teil automatisiert. Es verbindet eine Sequenzdatenbank mit einem Primerdesignprogramm und erreicht dadurch eine schnellere und robustere Auswahl geeigneter Primersequenzen.

Mit den vorgestellten Anwendungen wurde ein System geschaffen, welches als Datenbank- und Analyse-Netzwerk den experimentellen Ablauf der biomedizinischen Hochdurchsatzforschung ermöglicht. Es erlaubte die Untersuchung der Pathomechanismen unterschiedlicher Hirntumore (Menigiome, Medulloblastome) und hämatopoetischer Fragestellungen (Differenzierung der Zelllinie HL60, Akute Myeloische Leukämie). Insbesondere konnte die Entstehung und Progression des *Non-Melanom* Hautkrebses mit Hilfe von Microarrays mit murinen Fragmenten intensiv analysiert werden. Die Identifikation von neuen Kandidatengenomen für humanen Hautkrebs, unter anderem Gene des Epidermalen Differenzierungskomplexes (1q21.3), wurde dadurch ermöglicht.

Summary

The application of high-throughput technologies like *matrixCGH* and *expression profiling* with microarrays in biomedical research, allows for the highly parallel analysis of clinical samples. In order to conduct these kinds of experiments and cope with the amount of data produced, specialized software solutions are needed.

For the management of a clone collection, which is the starting point for all experiments, the *CloneBase* system was developed. It comprises of a database with internet access pages and automated data update functions. It permits the archiving, annotation and information retrieval for all clones. The nucleic acid fragments of selected clones are used for the production of microarrays with the help of robots. To operate and control the complex production process, a laboratory information management system (LIMS) was developed. *QuickLIMS* can process the barcode identified microtiter plates containing the fragments, using a specified protocol. The user interface is generated dynamically from the workflow, which itself is defined in a flexible manner. The system allows a maximum of process automation and error tracking for microarray production. The integration of bio-molecular data from diverse sources like public databases or clinical knowledge (annotation data) is a key challenge in the process of the analysis of these high-throughput experiments. The *Flexible Annotation and Correlation Tool (FACT)* was developed as a flexible framework for the explorative meta-analysis of genomic, proteomic or other experiments. Heterogeneous experimental data and annotational sources can be integrated and diverse analytical algorithms can be applied with the goal of finding distinct patterns and illuminating inherent characteristics within the data. The results from different experimental systems, e.g. information on genomic changes from *matrixCGH* analysis and expression profiles generated by cDNA microarrays can be brought into correlation using individual program modules. This is achieved by the abstraction from specific data types into a general data scheme and by the flexible storage of data source definitions. The *FACT* system is easily extendible to cover additional sources and include new methods. Finally, the application *AutoPrime* automates the process of selecting primer sequences for RQ-PCR as used for the verification of the achieved microarray results. It links a sequence database with a primer design program, making the selection faster and less error-prone.

With the applications introduced here, a database and analysis system was established, which facilitates the experimental process of biomedical high-throughput research. It allowed the successful analysis of pathomechanisms of different brain tumors (meningioma and medulloblastoma) and of hematopoietic topics (differentiation of the cell line HL60, acute myeloid leukemia). In particular the development and progression of *non-melanoma* skin cancer was comprehensively analyzed with microarrays containing murin fragments. This allowed the identification of new candidate genes for human skin cancer, among others members of the epidermal differentiation complex (1q21.3).