Open access publishing

A researcher's perspective

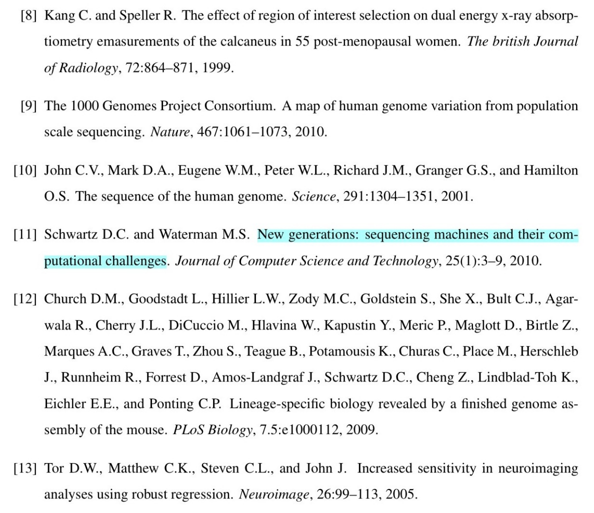
Karl Broman

Biostatistics & Medical Informatics
University of Wisconsin–Madison

`kbroman.org`
`github.com/kbroman`
`@kwbroman`

Access in action

Interesting reference



[8] Kang C. and Speller R. The effect of region of interest selection on dual energy x-ray absorptiometry emasurements of the calcaneus in 55 post-menopausal women. *The british Journal of Radiology*, 72:864–871, 1999.

[9] The 1000 Genomes Project Consortium. A map of human genome variation from population scale sequencing. *Nature*, 467:1061–1073, 2010.

[10] John C.V., Mark D.A., Eugene W.M., Peter W.L., Richard J.M., Granger G.S., and Hamilton O.S. The sequence of the human genome. *Science*, 291:1304–1351, 2001.

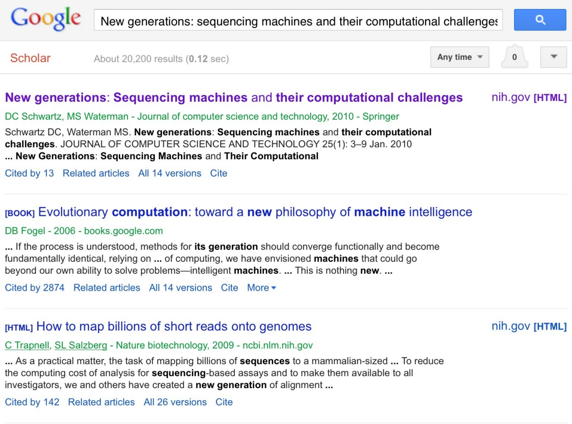
[11] Schwartz D.C. and Waterman M.S. New generations: sequencing machines and their computational challenges. *Journal of Computer Science and Technology*, 25(1):3–9, 2010.

[12] Church D.M., Goodstadt L., Hillier L.W., Zody M.C., Goldstein S., She X., Bult C.J., Agarwala R., Cherry J.L., DiCuccio M., Hlavina W., Kapustin Y., Meric P., Maglott D., Birtle Z., Marques A.C., Graves T., Zhou S., Teague B., Potamousis K., Churas C., Place M., Herschleb J., Runnheim R., Forrest D., Amos-Landgraf J., Schwartz D.C., Cheng Z., Lindblad-Toh K., Eichler E.E., and Ponting C.P. Lineage-specific biology revealed by a finished genome assembly of the mouse. *PLoS Biology*, 7.5:e1000112, 2009.

[13] Tor D.W., Matthew C.K., Steven C.L., and John J. Increased sensitivity in neuroimaging analyses using robust regression. *Neuroimage*, 26:99–113, 2005.

Access in action

Google Scholar



New generations: sequencing machines and their computational challenges

Scholar

About 20,200 results (**0.12** sec)

New generations: **Sequencing machines** and **their computational challenges** nih.gov **[HTML]**

DC Schwartz, MS Waterman - Journal of computer science and technology, 2010 - Springer

Schwartz DC, Waterman MS. **New generations**: **Sequencing machines** and **their computational challenges**. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 25(1): 3–9 Jan. 2010 ... **New Generations**: **Sequencing Machines** and **Their Computational**

Cited by 13 Related articles All 14 versions Cite

[BOOK] Evolutionary **computation**: toward a **new** philosophy of **machine** intelligence

DB Fogel - 2006 - books.google.com

... If the process is understood, methods for **its generation** should converge functionally and become fundamentally identical, relying on ... of computing, we have envisioned **machines** that could go beyond our own ability to solve problems—intelligent **machines**. ... This is nothing **new**. ...

Cited by 2874 Related articles All 14 versions Cite More ▾

[HTML] How to map billions of short reads onto genomes nih.gov **[HTML]**

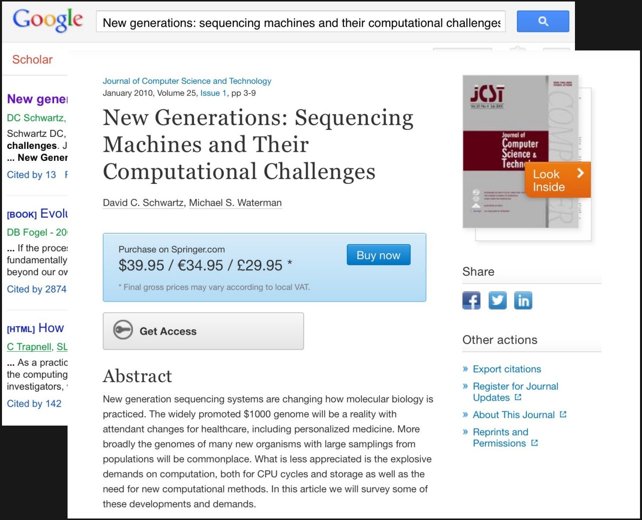
C Trapnell, SL Salzberg - Nature biotechnology, 2009 - ncbi.nlm.nih.gov

... As a practical matter, the task of mapping billions of **sequences** to a mammalian-sized ... To reduce the computing cost of analysis for **sequencing**-based assays and to make them available to all investigators, we and others have created a **new generation** of alignment ...

Cited by 142 Related articles All 26 versions Cite

Access in action

Google Scholar



New generations: sequencing machines and their computational challenges

New gene
DC Schwartz,
Schwartz DC,
challenges. J
... **New Gener**
Cited by 13 F

[BOOK] Evolu
DB Fogel - 20
... If the proce
fundamentally
beyond our ov
Cited by 2874

[HTML] How
C Trapnell, SL
... As a practic
the computing
investigators,
Cited by 142

Journal of Computer Science and Technology
January 2010, Volume 25, Issue 1, pp 3-9

New Generations: Sequencing Machines and Their Computational Challenges

David C. Schwartz, Michael S. Waterman

Purchase on Springer.com
$39.95 / €34.95 / £29.95 *
* Final gross prices may vary according to local VAT.

Abstract

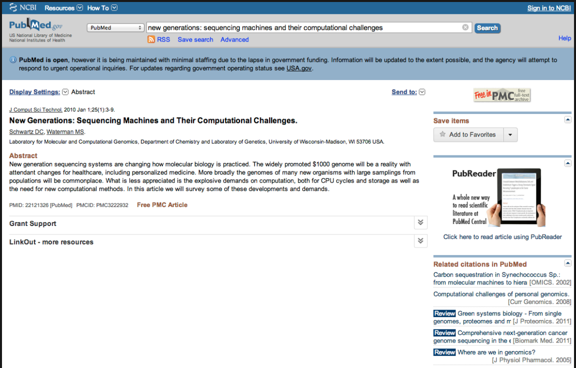
New generation sequencing systems are changing how molecular biology is practiced. The widely promoted $1000 genome will be a reality with attendant changes for healthcare, including personalized medicine. More broadly the genomes of many new organisms with large samplings from populations will be commonplace. What is less appreciated is the explosive demands on computation, both for CPU cycles and storage as well as the need for new computational methods. In this article we will survey some of these developments and demands.

What's the deal with the prices?

Broman K, Speed T, Tigges M (1996) Estimation of antigen-responsive T cell frequencies in PBMC from human subjects. J Immunol Meth 198:119–132	$39.95
Broman KW, Weber JL (1999) Method for constructing confidently ordered linkage maps. Genet Epidemiol 16:337–343	$35.00
Broman KW, Feingold E (2004) SNPs made routine. Nat Methods 1:104–105	$18.00
Broman KW (2005) Mapping expression in randomized rodent genomes. Nat Genet 37:209–210	$18.00

Access in action

There's also PubMed



PubMed is open, however it is being maintained with minimal staffing due to the lapse in government funding. Information will be updated to the extent possible, and the agency will attempt to respond to urgent operational inquiries. For updates regarding government operating status see USA.gov.

J Comput Sci Technol. 2010 Jan 1;25(1):3-9.

New Generations: Sequencing Machines and Their Computational Challenges.

Schwartz DC, Waterman MS.

Laboratory for Molecular and Computational Genomics, Department of Chemistry and Laboratory of Genetics, University of Wisconsin-Madison, WI 53706 USA.

Abstract

New generation sequencing systems are changing how molecular biology is practiced. The widely promoted $1000 genome will be a reality with attendant changes for healthcare, including personalized medicine. More broadly the genomes of many new organisms with large samplings from populations will be commonplace. What is less appreciated is the explosive demands on computation, both for CPU cycles and storage as well as the need for new computational methods. In this article we will survey some of these developments and demands.

PMID: 22121326 [PubMed] PMCID: PMC3222932 **Free PMC Article**

Related citations in PubMed

Carbon sequestration in Synechococcus Sp.: from molecular machines to hiera [OMICS. 2002]

Computational challenges of personal genomics. [Curr Genomics. 2008]

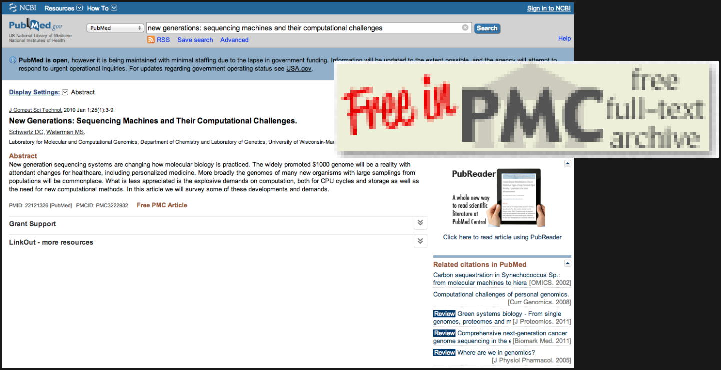
Review Green systems biology - From single genomes, proteomes and m [J Proteomics. 2011]

Review Comprehensive next-generation cancer genome sequencing in the ε [Biomark Med. 2011]

Review Where are we in genomics? [J Physiol Pharmacol. 2005]

Access in action

There's also PubMed



new generations: sequencing machines and their computational challenges

PubMed is open, however it is being maintained with minimal staffing due to the lapse in government funding. Information will be updated to the extent possible, and the agency will attempt to respond to urgent operational inquiries. For updates regarding government operating status see USA.gov.

Display Settings: Abstract

J Comput Sci Technol. 2010 Jan 1;25(1):3-9.

New Generations: Sequencing Machines and Their Computational Challenges.

Schwartz DC, Waterman MS.

Laboratory for Molecular and Computational Genomics, Department of Chemistry and Laboratory of Genetics, University of Wisconsin-Mad

Abstract

New generation sequencing systems are changing how molecular biology is practiced. The widely promoted $1000 genome will be a reality with attendant changes for healthcare, including personalized medicine. More broadly the genomes of many new organisms with large samplings from populations will be commonplace. What is less appreciated is the explosive demands on computation, both for CPU cycles and storage as well as the need for new computational methods. In this article we will survey some of these developments and demands.

PMID: 22121326 [PubMed] PMCID: PMC3222932 **Free PMC Article**

Grant Support

LinkOut - more resources

Click here to read article using PubReader

Related citations in PubMed

- Carbon sequestration in Synechococcus Sp.: from molecular machines to hiera [OMICS. 2002]
- Computational challenges of personal genomics. [Curr Genomics. 2008]
- Review Green systems biology - From single genomes, proteomes and m [J Proteomics. 2011]
- Review Comprehensive next-generation cancer genome sequencing in the e [Biomark Med. 2011]
- Review Where are we in genomics? [J Physiol Pharmacol. 2005]

Another example

PubMed is open, however it is being maintained with minimal staffing due to the lapse in government funding. Information will be updated to the extent possible, and the agency will attempt to respond to urgent operational inquiries. For updates regarding government operating status see USA.gov.

J Dent Res. 2013 Jan;92(1):32-7. doi: 10.1177/0022034512463241. Epub 2012 Oct 11.

Clustering tooth surfaces into biologically informative caries outcomes.

Shaffer JR, Feingold E, Wang X, Weeks DE, Weyant RJ, Crout R, McNeil DW, Marazita ML.

Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA. jr8@pitt.edu

Abstract

Dental caries affects most adults worldwide; however, the risk factors for dental caries do not necessarily exert their effects uniformly across all tooth surfaces. Instead, the actions of some risk factors may be limited to a subset of teeth/surfaces. Therefore, we used hierarchical clustering on tooth surface-level caries data for 1,068 Appalachian adults (ages 18-75 yrs) to group surfaces based on co-occurrence of caries. Our cluster analysis yielded evidence of 5 distinct groups of tooth surfaces that differ with respect to caries: (C1) pit and fissure molar surfaces, (C2) mandibular anterior surfaces, (C3) posterior non-pit and fissure surfaces, (C4) maxillary anterior surfaces, and (C5) mid-dentition surfaces. These clusters were replicated in a national dataset (NHANES 1999-2000, N = 3,123). We created new caries outcomes defined as the number of carious tooth surfaces within each cluster. We show that some cluster-based caries outcomes are heritable (i.e., under genetic regulation; $p < 0.05$), whereas others are not. Likewise, we demonstrate the association between some cluster-based caries outcomes and potential risk factors such as age, sex, educational attainment, and toothbrushing habits. Together, these results suggest that the permanent dentition can be subdivided into groups of tooth surfaces that are useful for understanding the factors influencing cariogenesis. Abbreviations: COHRA, Center for Oral Health in Appalachia, the principal study sample; C1-5, clusters 1-5, groups of similarly behaving tooth surfaces identified through hierarchical clustering; DMFS index, decayed, missing, or filled surfaces, a traditional caries measure representing the number of affected surfaces across the entire dentition; DMFS1-5, partial DMFS indices representing the number of affected surfaces within a hierarchical cluster; and NHANES, National Health and Nutrition Examination Survey, the secondary study sample.

PMID: 23064960 [PubMed - indexed for MEDLINE] PMCID: PMC3521447 [Available on 2014/1/1]

Save items

Related citations in PubMed

- Genetic susceptibility to dental caries on pit and fissure and sr [Caries Res. 2012]
- GWAS of dental caries patterns in the permanent dentition. [J Dent Res. 2013]
- Incremental susceptibility of individual [Community Dent Oral Epidemiol. 1996]
- Comparison of tooth surface-specific dental caries attack p [J Dent Res. 1993]
- **Review** [Factors that modify de- and remineraliz [Ann Acad Med Stetin. 1999]

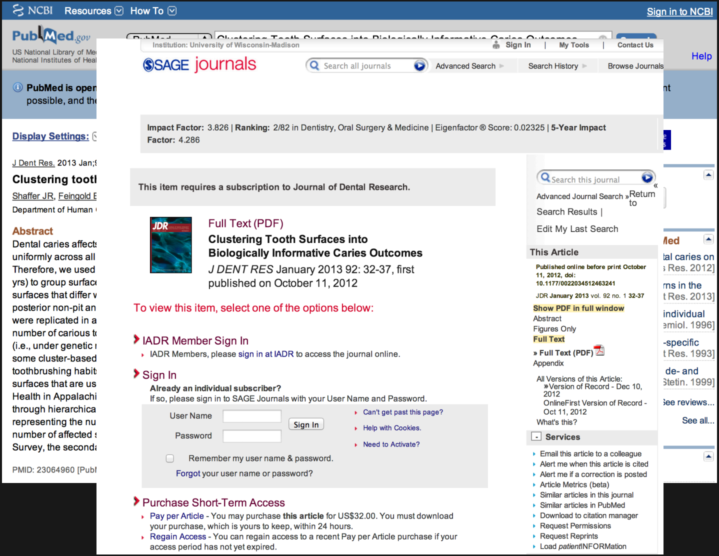
See reviews...

See all...

Related information

Related Citations

Another example



J Dent Res. 2013 Jan;

Clustering tooth

Shaffer JR, Feingold E

Department of Human

Abstract

Dental caries affects
uniformly across all
Therefore, we used
yrs) to group surfac
surfaces that differ v
posterior non-pit an
were replicated in a
number of carious t
(i.e., under genetic i
some cluster-based
toothbrushing habit
surfaces that are us
Health in Appalachi
through hierarchica
representing the nu
number of affected s
Survey, the second

PMID: 23064960 [Publ

Impact Factor: 3.826 | **Ranking:** 2/82 in Dentistry, Oral Surgery & Medicine | Eigenfactor ® Score: 0.02325 | **5-Year Impact Factor:** 4.286

This item requires a subscription to Journal of Dental Research.

Full Text (PDF)

Clustering Tooth Surfaces into Biologically Informative Caries Outcomes

J DENT RES January 2013 92: 32-37, first published on October 11, 2012

To view this item, select one of the options below:

IADR Member Sign In

- IADR Members, please sign in at IADR to access the journal online.

Sign In

Already an individual subscriber?
If so, please sign in to SAGE Journals with your User Name and Password.

User Name
Password
Sign In
☐ Remember my user name & password.
Forgot your user name or password?

- Can't get past this page?
- Help with Cookies.
- Need to Activate?

Purchase Short-Term Access

- Pay per Article - You may purchase **this article** for US$32.00. You must download your purchase, which is yours to keep, within 24 hours.
- Regain Access - You can regain access to a recent Pay per Article purchase if your access period has not yet expired.

This Article

Published online before print October 11, 2012, doi: 10.1177/0022034512463241

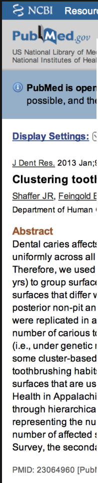
JDR **January 2013** vol. 92 no. 1 **32-37**

All Versions of this Article:
»Version of Record - Dec 10, 2012
OnlineFirst Version of Record - Oct 11, 2012
What's this?

Services

- Email this article to a colleague
- Alert me when this article is cited
- Alert me if a correction is posted
- Article Metrics (beta)
- Similar articles in this journal
- Similar articles in PubMed
- Download to citation manager
- Request Permissions
- Request Reprints
- Load *patient*INFORMation

Another example



US National Library of Me
National Institutes of Heal

PubMed is open
possible, and th

Display Settings:

J Dent Res. 2013 Jan;

Clustering tootl

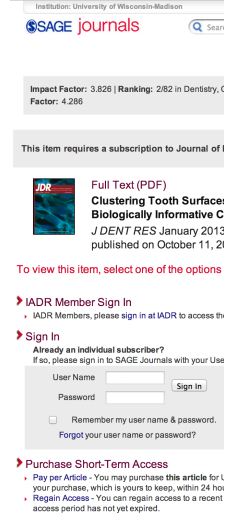
Shaffer JR, Feingold E

Department of Human

Abstract

Dental caries affect
uniformly across all
Therefore, we used
yrs) to group surfac
surfaces that differ v
posterior non-pit an
were replicated in a
number of carious t
(i.e., under genetic r
some cluster-based
toothbrushing habit
surfaces that are us
Health in Appalachi
through hierarchica
representing the nu
number of affected s
Survey, the second

PMID: 23064960 [Pub



Impact Factor: 3.826 | **Ranking:** 2/82 in Dentistry, C
Factor: 4.286

This item requires a subscription to Journal of I

Full Text (PDF)

Clustering Tooth Surface
Biologically Informative C

J DENT RES January 2013
published on October 11, 20

To view this item, select one of the options

IADR Member Sign In

- IADR Members, please sign in at IADR to access th

Sign In

Already an individual subscriber?
If so, please sign in to SAGE Journals with your Use

User Name
Password
Sign In

☐ Remember my user name & password.

Forgot your user name or password?

Purchase Short-Term Access

- Pay per Article - You may purchase **this article** for U
your purchase, which is yours to keep, within 24 hou
- Regain Access - You can regain access to a recent
access period has not yet expired.

RESEARCH REPORTS
Clinical

J.R. Shaffer[1,2*], E. Feingold[1,3],
X. Wang[2,4,5], D.E. Weeks[1,3],
R.J. Weyant[2,6], R. Crout[2,7],
D.W. McNeil[8], and M.L. Marazita[1,2,4,5,9]

[1]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA; [2]Center for Oral Health Research in Appalachia, University of Pittsburgh, Pittsburgh, PA 15261, USA, and West Virginia University, Morgantown, WV 26506, USA; [3]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA; [4]Center for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA 15219, USA; [5]Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA; [6]Department of Dental Public Health and Information Management, University of Pittsburgh, School of Dental Medicine, Pittsburgh, PA 15261, USA; [7]Department of Periodontics, West Virginia University School of Dentistry, Morgantown, WV 26506, USA; [8]Dental Practice and Rural Health, West Virginia University School of Dentistry, Morgantown, WV 26506, USA; and [9]Clinical and Translational Science Institute, and Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; *corresponding author, jrs51@pitt.edu

J Dent Res 92(1):32-37, 2013

Clustering Tooth Surfaces into Biologically Informative Caries Outcomes

ABSTRACT

Dental caries affects most adults worldwide; however, the risk factors for dental caries do not necessarily exert their effects uniformly across all tooth surfaces. Instead, the actions of some risk factors may be limited to a subset of teeth/surfaces. Therefore, we used hierarchical clustering on tooth surface-level caries data for 1,068 Appalachian adults (ages 18-75 yrs) to group surfaces based on co-occurrence of caries. Our cluster analysis yielded evidence of 5 distinct groups of tooth surfaces that differ with respect to caries: (C1) pit and fissure molar surfaces, (C2) mandibular anterior surfaces, (C3) posterior non-pit and fissure surfaces, (C4) maxillary anterior surfaces, and (C5) mid-dentition surfaces. These clusters were replicated in a national dataset (NHANES 1999-2000, N = 3,123). We created new caries outcomes defined as the number of carious tooth surfaces within each cluster. We show that some cluster-based caries outcomes are heritable (*i.e.*, under genetic regulation; $p < 0.05$), whereas others are not. Likewise, we demonstrate the association between some cluster-based caries outcomes and potential risk factors such as age, sex, educational attainment, and toothbrushing habits. Together, these results suggest that the permanent dentition can be subdivided into groups of tooth surfaces that are useful for understanding the factors influencing cariogenesis. *Abbreviations*: COHRA, Center for Oral Health in Appalachia, the principal study sample; C1-5, clusters 1-5, groups of similarly behaving tooth surfaces identified through hierarchical clustering; DMFS index, decayed, missing, or filled surfaces, a traditional caries measure representing the number of affected surfaces across the entire dentition; DMFS1-5, partial DMFS indices representing the number of affected surfaces within a hierarchical cluster; and NHANES, National Health and Nutrition Examination Survey, the secondary study sample.

KEY WORDS: dental caries, permanent dentition, white spots, hierarchical clustering, cluster analysis, heritability.

DOI: 10.1177/0022034512463241

Received June 29, 2012; Last revision August 30, 2012; Accepted September 8, 2012

A supplemental appendix to this article is published electronically only at http://jdr.sagepub.com/supplemental.

© International & American Associations for Dental Research

32

INTRODUCTION

Dental caries, which affects the great majority of adolescents and adults throughout the world, is a multi-factorial disease caused by the effects of numerous environmental, behavioral, and genetic factors. Many risk factors have been identified, such as: host genetics (Horowitz *et al.*, 1958); environmental exposures, including fluoride, cariogenic bacteria, and pH-altering agents; behavioral factors, including diet and oral hygiene; characteristics of the dentition, including enamel composition and positions and morphology of teeth; characteristics of the oral environment, including saliva composition, flow rate, and pH buffering capacity; and demographic factors, including age, sex, race, ethnicity, socio-economic status, and access to oral health care (Hunter, 1988). This complexity is further compounded by the disease phenotype, which may manifest as innumerable combinations of caries lesions across tooth surfaces of the full dentition. Because caries risk factors may exert differential effects across tooth surfaces of the permanent dentition, measurable caries experience may be modeled as the cumulative result of multiple superimposed patterns of decay due to the various risk factors (Batchelor and Sheiham, 2004; Shaffer *et al.*, 2012a).

In epidemiological studies, however, caries experience is typically reduced to a single measure of decay, such as the widely used DMFT/S index (calculated as the sum of decayed, missing due to decay, or filled/restored teeth/surfaces). Such global measures of caries experience ignore the fact that categories of tooth surfaces exhibit differences in susceptibility to decay and are differentially affected by risk factors. Because they ignore the patterns of dental caries across the dentition, global measures of decay such as DMFT/S index may be limited in their ability to identify caries risk factors that exert their effects on specific categories of tooth surfaces. Indeed, previous studies have demonstrated that modeling the patterns of tooth decay is beneficial for epidemiological (Psoter *et al.*, 2003, 2009) and genetic studies (Shaffer *et al.*, 2012a, 2012b). We used hierarchical clustering analysis to group tooth surfaces into categories based on co-occurrence of caries lesions. We then generated novel caries outcomes reflecting these tooth-surface categories, and explored their utility for studying caries etiology.

Another example

PubMed is open
possible, and the

Display Settings:

J Dent Res. 2013 Jan;9

Clustering tootl

Shaffer JR, Feingold E

Department of Human

Abstract

Dental caries affects
uniformly across all
Therefore, we used
yrs) to group surfac
surfaces that differ v
posterior non-pit an
were replicated in a
number of carious t
(i.e., under genetic r
some cluster-based
toothbrushing habit
surfaces that are us
Health in Appalachi
through hierarchica
representing the nu
number of affected s
Survey, the second

PMID: 23064960 [Publ

Impact Factor: 3.826 | **Ranking:** 2/82 in Dentistry, C
Factor: 4.286

This item requires a subscription to Journal of I

Full Text (PDF)

Clustering Tooth Surfaces
Biologically Informative C

J DENT RES January 2013
published on October 11, 20

To view this item, select one of the options

IADR Member Sign In

IADR Members, please sign in at IADR to acce

Sign In

Already an individual subscriber?
If so, please sign in to SAGE Journals with you

User Name
Password
Sign In
☐ Remember my user name & passw
Forgot your user name or password?

Purchase Short-Term Access

- Pay per Article - You may purchase **this articl** your purchase, which is yours to keep, within 2
- Regain Access - You can regain access to a r access period has not yet expired.

RESEARCH REPORTS
Clinical

J.R. Shaffer[1,2*], E. Feingold[1,3],
X. Wang[2,4,5], D.E. Weeks[1,3],
R.J. Weyant[2,6], R. Crout[2,7],
D.W. McNeil[8], and M.L. Marazita[1,2,4,5,9]

[1]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA; [2]Center for Oral Health Research in Appalachia, University of Pittsburgh, Pittsburgh, PA 15261, USA, and West Virginia University, Morgantown, WV 26506, USA; [3]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA; [4]Center for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA 15219, USA; [5]Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA; [6]Department of Dental Public Health and Information Management, University of Pittsburgh, School of Dental Medicine, Pittsburgh, PA 15261, USA; [7]Department of Periodontics, West Virginia University School of Dentistry, Morgantown, WV 26506, USA; [8]Dental Practice and Rural Health, West Virginia University School of Dentistry, Morgantown, WV 26506, USA; and [9]Clinical and Translational Science Institute, and Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; *corresponding author, jrs51@pitt.edu

J Dent Res 92(1):32-37, 2013

Clustering Tooth Surfaces into Biologically Informative Caries Outcomes

ABSTRACT

Dental caries affects most adults worldwide; however, the risk factors for dental caries do not necessarily exert their effects uniformly across all tooth surfaces. Instead, the actions of some risk factors may be limited to a subset of teeth/surfaces. Therefore, we used hierarchical clustering on tooth surface-level caries data for 1,068 Appalachian adults (ages 18-75 yrs) to group surfaces based on co-occurrence of caries. Our cluster analysis yielded evidence of 5 distinct groups of tooth surfaces that differ with respect to caries: (C1) pit and fissure molar surfaces, (C2) mandibular anterior surfaces, (C3) posterior non-pit and fissure surfaces, (C4) maxillary anterior surfaces, and (C5) mid-dentition surfaces. These clusters were replicated in a national dataset (NHANES 1999-2000, N = 3,123). We created new caries outcomes defined as the number of carious tooth surfaces within each cluster. We show that some cluster-based caries outcomes are heritable (*i.e.*, under genetic regulation; $p < 0.05$), whereas others are not. Likewise, we demonstrate the association between some cluster-based

INTRODUCTION

Dental caries, which affects the great majority of adolescents and adults throughout the world, is a multi-factorial disease caused by the effects of numerous environmental, behavioral, and genetic factors. Many risk factors have been identified, such as: host genetics (Horowitz *et al*., 1958); environmental exposures, including fluoride, cariogenic bacteria, and pH-altering agents; behavioral factors, including diet and oral hygiene; characteristics of the dentition, including enamel composition and positions and morphology of teeth; characteristics of the oral environment, including saliva composition, flow rate, and pH buffering capacity; and demographic factors, including age, sex, race, ethnicity, socio-economic status, and access to oral health care (Hunter, 1988). This complexity is further compounded by the disease phe- ries lesions
factors may
tition, mea-
of multiple
tchelor and

lly reduced
dex (calcu-
tored teeth/
act that cat-
cay and are
patterns of
as DMFT/S
s that exert
ious studies
neficial for
haffer *et al*.,
p tooth sur-
then gener-
gories, and

DOI: 10.1177/0022034512463241

Received June 29, 2012; Last revision August 30, 2012; Accepted September 8, 2012

A supplemental appendix to this article is published electronically only at http://jdr.sagepub.com/supplemental.

32

Another example

NCBI Resources How To

PubMed.gov

US National Library of Me
National Institutes of Heal

PubMed is open
possible, and the

Display Settings:

J Dent Res. 2013 Jan;9

Clustering tootl

Shaffer JR, Feingold E

Department of Human

Abstract

Dental caries affect
uniformly across all
Therefore, we used
yrs) to group surfac
surfaces that differ v
posterior non-pit an
were replicated in a
number of carious t
(i.e., under genetic i
some cluster-based
toothbrushing habit
surfaces that are us
Health in Appalachi
through hierarchica
representing the nu
number of affected s
Survey, the second

PMID: 23064960 [Publ

Institution: University of Wisconsin-Madison

SAGE journals

Searc

Impact Factor: 3.826 | **Ranking:** 2/82 in Dentistry, C
Factor: 4.286

This item requires a subscription to Journal of

Full Text (PDF)

Clustering Tooth Surfaces
Biologically Informative C

J DENT RES January 2013
published on October 11, 20

To view this item, select one of the options

IADR Member Sign In

- IADR Members, please sign in at IADR to acce

Sign In

Already an individual subscriber?
If so, please sign in to SAGE Journals with you

User Name
Password
Sign In

☐ Remember my user name & passw

Forgot your user name or password?

Purchase Short-Term Access

- Pay per Article - You may purchase this articl
your purchase, which is yours to keep, within 2
- Regain Access - You can regain access to a r
access period has not yet expired.

RESEARCH REPORTS
Clinical

J.R. Shaffer[1,2]
X. Wang[2,4,5],
R.J. Weyant[2],
D.W. McNeil

J Dent Res 92(1):

ABSTRACT

DOI: 10.1

Received
Accepted

A supplem
tronically

© Internati

32

Institution: University of Wisconsin-Madison | Sign In | My Tools | Conta

SAGE journals | Search all journals | Advanced Search | Search History | Browse

JDR JOURNAL OF DENTAL RESEARCH®

Home | OnlineFirst | All Issues | Subscribe | RSS | Email Alerts

« Return to Search Results | Edit My Last Search | Search this journal

Impact Factor: 3.826 | **Ranking:** 2/82 in Dentistry, Oral Surgery & Medicine | Eigenfactor® Score: 0.02325 | **5-Year Impact Factor:** 4.286

This item requires a subscription to Journal of Dental Research.

IADR Member Sign In

- IADR Members, please sign in at IADR to access the journal online.

Sign In

Already an individual subscriber?
If so, please sign in to SAGE Journals with your User Name and Password.

User Name
Password
Sign In

- Can't get past this page?
- Help with Cookies.
- Need to Activate?

☐ Remember my user name & password.

Forgot your user name or password?

OpenAthens Users

- Sign in via OpenAthens : If your organization uses OpenAthens, you can log in using your OpenAthens username and password. Contact your library for more details.
- List of OpenAthens registered sites, including contact details.

Login via Your Institution

- Login via your institution : You may be able to gain access using your login credentials for your institution. Contact your library if you do not have a username and password.

Subscribe/Recommend

- Become a Member - Join the International and American Associations for Dental Research.
- Click here to subscribe to the print and/or online journal.
- Click here to recommend to your library.

Current Issue

- October 2013, 92 (10)
- Alert me to new issues of Journal of Dental Research

Submit a Manuscript
Free Sample Copy
Email Alerts
RSS feed

More about this journal

Other IADR Journals

Advances in Dental Research
Critical Reviews in Oral Biology & Medicine

- International and American Associations for Dental Research

Interested in article-level

Twitter is useful

(for venting)

Karl Broman
@kwbroman

You SAGE ba[redacted]rds. ILL'd an article from J Dent Res, but it didn't include supplement, also behind pay wall. jdr.sagepub.com/content/92/1/3... #OA

10/14/13, 11:23 AM

Twitter is useful

(for venting and more)

Karl Br
@kwbr

You SAGE b
from J Dent
supplement,
jdr.sagepub.
#OA

10/14/13, 11:23 A

RESEARCH REPORTS
Clinical

J.R. Shaffer[1,2*], E. Feingold[1,3], X. Wang[2,4,5],
D.E. Weeks[1,3], R.J. Weyant[2,6], R. Crout[2,7],
D.W. McNeil[8], and M.L. Marazita[1,2,4,5,9]

[1]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA; [2]Center for Oral Health Research in Appalachia, University of Pittsburgh, Pittsburgh, PA 15261, USA, and West Virginia University, Morgantown, WV 26506, USA; [3]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA; [4]Center for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA 15219, USA; [5]Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA; [6]Department of Dental Public Health and Information Management, University of Pittsburgh, School of Dental Medicine, Pittsburgh, PA 15261, USA; [7]Department of Periodontics, West Virginia University School of Dentistry, Morgantown, WV 26506, USA; [8]Dental Practice and Rural Health, West Virginia University School of Dentistry, Morgantown, WV 26506, USA; and [9]Clinical and Translational Science Institute, and Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; *corresponding author, jrs51@pitt.edu

J Dent Res DOI: 10.1177/0022034512463241

Clustering Tooth Surfaces into Biologically Informative Caries Outcomes

APPENDIX

COHRA Population Characteristics

Descriptive characteristics of the COHRA sample are shown in the Appendix Table. Compared with the general US population, this Appalachian sample from rural West Virginia and Pennsylvania is poorer, less educated, and underserved with respect to oral health care.

Cross-validation and Defining Separate Clusters

Determining the number of separate clusters is the principal challenge in the field of cluster analysis, and there is currently very little statistical theory in this area. Defining *a priori* minimum distances between clusters is arbitrary and may lead to over-fitting. Therefore, we instead performed two-fold cross-validation (*e.g.,* Salvador and Chan, 2004), a commonly used approach for determining the number of clusters. Cross-validation allows us to determine the sensitivity of our clusters to perturbations of the input data and to identify what level of within-cluster similarity and between-cluster dissimilarity defines stable clusters. We randomly divided our sample into two halves and performed hierarchical clustering on each half. This process was repeated for 10 random halves. By comparing dendrograms, we determined the maximum number of separate clusters that were consistently observed across all random halves. Overall, tooth surfaces were distributed across 5 very stable clusters. Example cluster results for two complementary halves (which total to the full sample) are shown in Appendix Fig. 1. Similarly, to assess any effects of the inclusion of biological relatives in our sample, we repeated hierarchical clustering in the maximal subset of unrelated individuals (see Appendix Fig. 2), which were nearly identical to the total COHRA and NHANES 1999-2000 samples.

In both the total COHRA sample and in the NHANES 1999-2000 sample, the fifth cluster was subdivided into maxillary and mandibular components (indicating possibly 6 instead of 5 clusters). However, these subdivisions were not consistently observed *via* cross-validation of the COHRA sample. That is, in some randomly chosen halves of the COHRA sample, the sixth cluster did not distinguish maxillary and mandibular components of the fifth cluster. Therefore, we have conservatively shown results for the 5 stable clusters, as well as the maxillary and mandibular subdivisions of the fifth cluster.

Overwhelmingly, the cluster results were stable within the COHRA sample, and consistent between the COHRA sample and NHANES 1999-2000 sample. That being stated, there were subtle differences observed among cluster results. For example, left-right asymmetry was observed, albeit rarely, in the random halves (*e.g.,* Appendix Fig. 1A, tooth #20 and tooth #29). Likewise, some tooth surfaces physically positioned on the border between two adjacent clusters shifted membership (*e.g.,* Appendix Figs. 1A and 2, tooth #21 and tooth #28). In many cases, the result of such shifts in cluster membership echoed the subtle differences between COHRA and NHANES 1999-2000. We speculate that the all-or-nothing nature of our clustering approach represents an oversimplification of the relationships among tooth surfaces with respect to dental caries risk factors

Twitter is useful

#icanhazpdf

It's all about money

(Costs in scientific publishing)

- Research
- Writing
- Peer review, editorial oversight
- Journal administration
- Copy editing, typesetting
- Distribution

It's all about money

(Costs in scientific publishing)

- Research
- Writing
- Peer review, editorial oversight
- Journal administration
- Copy editing, typesetting
- Distribution
- Profit

It's all about money

(Costs in scientific publishing)

- Research
- Writing
- Peer review, editorial oversight
- Journal administration
- Copy editing, typesetting
- Distribution
- Profit

It's all about money

(Costs in scientific publishing)

- Research
- Writing
- Peer review, editorial oversight
- Journal administration
- Copy editing, typesetting
- Distribution
- Profit

It's not about

- Peer review
- Predatory publishing
- Impact factors
- Evaluating researchers
  (for grants & promotions)

It's not about

- Peer review
- Predatory publishing
- Impact factors
- Evaluating researchers
  (for grants & promotions)

Well, it sort of is…

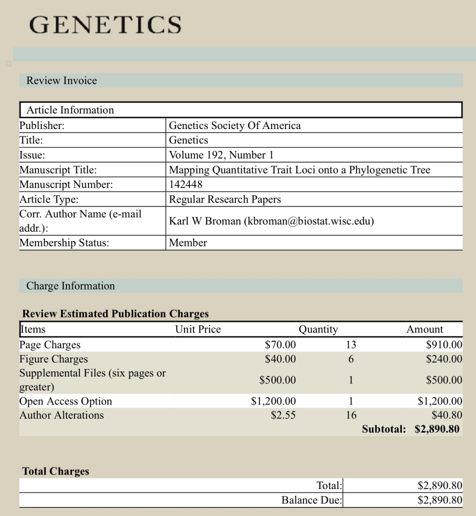
Paying for it

- Traditional approach
  - subscriptions
  - page charges
- Open access
  - bigger page charges
  - submission charges?

Paying for it

- Traditional approach
  - subscriptions
  - page charges
- Open access
  - bigger page charges
  - submission charges?
- Endowments
- Direct grants to journals

Invoice



GENETICS

Review Invoice

Article Information

Publisher:	Genetics Society Of America
Title:	Genetics
Issue:	Volume 192, Number 1
Manuscript Title:	Mapping Quantitative Trait Loci onto a Phylogenetic Tree
Manuscript Number:	142448
Article Type:	Regular Research Papers
Corr. Author Name (e-mail addr.):	Karl W Broman (kbroman@biostat.wisc.edu)
Membership Status:	Member

Charge Information

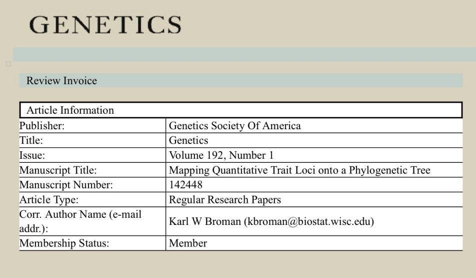
Review Estimated Publication Charges

Items	Unit Price	Quantity	Amount
Page Charges	$70.00	13	$910.00
Figure Charges	$40.00	6	$240.00
Supplemental Files (six pages or greater)	$500.00	1	$500.00
Open Access Option	$1,200.00	1	$1,200.00
Author Alterations	$2.55	16	$40.80
		Subtotal:	**$2,890.80**

Total Charges

Total:	$2,890.80
Balance Due:	$2,890.80

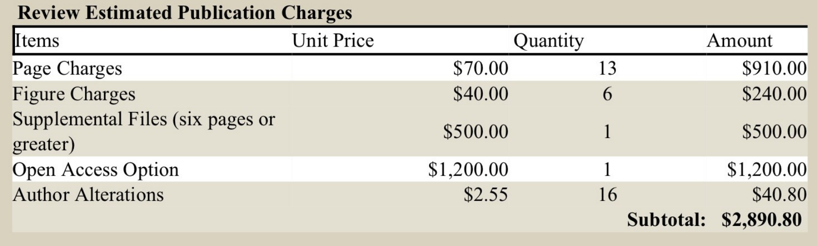
Invoice



GENETICS

Review Invoice

Article Information	
Publisher:	Genetics Society Of America
Title:	Genetics
Issue:	Volume 192, Number 1
Manuscript Title:	Mapping Quantitative Trait Loci onto a Phylogenetic Tree
Manuscript Number:	142448
Article Type:	Regular Research Papers
Corr. Author Name (e-mail addr.):	Karl W Broman (kbroman@biostat.wisc.edu)
Membership Status:	Member



Review Estimated Publication Charges

Items	Unit Price	Quantity	Amount
Page Charges	$70.00	13	$910.00
Figure Charges	$40.00	6	$240.00
Supplemental Files (six pages or greater)	$500.00	1	$500.00
Open Access Option	$1,200.00	1	$1,200.00
Author Alterations	$2.55	16	$40.80
		Subtotal:	**$2,890.80**

Choices for young investigators

- Pay for open access
- Support young open access journals

OR

- Let subscribers pay & do more experiments
- Continue to go after Science, Nature, & Cell

What can we do?

- Send our best work to open access journals
- Support junior faculty to keep their papers open
- Pay attention to the quality of the work
  (not the impact factor of the journal)
- Raise endowments for trusted journals
- Reform copyright law

What can we do?

- Send our best work to open access journals
- Support junior faculty to keep their papers open
- Pay attention to the quality of the work
  (not the impact factor of the journal)
- Raise endowments for trusted journals
- Reform copyright law