

Fifteen million records and counting: BIOSIS maintains a tradition of high quality life sciences data

Cindi Fitzgerald
Thomson Scientific
September 2005

On one hand, the fifteen millionth record in the BIOSIS[®] database is simply an article about population genetics. On the other, it is a milestone that represents a significant and wider achievement — the long-established value, reliability, and reputation of BIOSIS data and products.

Introducing record number 15 million

The article “*Detect and adjust for population stratification in population-based association study using genomic control markers; An application of Affymetrix Genechip[®] Human Mapping 10K array*”, published in the *European Journal of Human Genetics*, is the fifteen millionth BIOSIS record to be indexed since 1969, when the electronic file began.

Based on a topic — mathematics and statistical analysis in biology — that has attained notable attention in recent years, this article exemplifies the continuing relevance and usefulness of BIOSIS to the life sciences research community. It also shows how the 2004 acquisition of BIOSIS by Thomson Scientific has expanded BIOSIS' content and accessibility.

A mutually beneficial partnership provides more data and options

Since both Thomson Scientific and BIOSIS have long been dedicated to providing high quality research tools and technologies, the integration of the BIOSIS family of products into Thomson Scientific has been fitting and seamless.

The Thomson Scientific research platform, *ISI Web of KnowledgeSM*, adds significant value to BIOSIS content (figure 1). It allows users to cross search records in other *ISI Web of Knowledge* products that an institution subscribes to, including *Web of Science* — where users can also find the fifteen millionth record (figure 2), as well as valuable citation data that reveals the record's past influences and future implications. This cross-search capability enables users to take advantage of the special attributes and capabilities of several products simultaneously as they explore their research topic.

Via *ISI Web of Knowledge*, users can also create citation alerts, which inform them when a record is cited — providing a way to track further developments. Users can also analyze their full research results; grouping records by fields specific to BIOSIS products and gaining further insight into research trends and patterns.

Figure 1: The fifteen millionth BIOSIS record, viewed on the *ISI Web of Knowledge* platform

ISI Web of KnowledgeSM BIOSIS Previews

BIOSIS Previews[®]

Full Record

◀ Record 4 of 9 (Set #1) ▶

Accession Number: PREV200500079985

Document Type: Article

Title: Detect and adjust for population stratification in population-based association study using genomic control markers: an application of Affymetrix Genechip(R) Human Mapping 10K array

Author(s): [Hao, Ke](#); [Li, Cheng](#); [Rosenow, Carsten](#); [Wong, Wing H.](#) (wwong@hsph.harvard.edu)

Source: European Journal of Human Genetics 12 (12) : 1001-1006 December 2004

Language: English **Medium:** print

Abstract: Population-based association design is often compromised by false or nonreplicable findings, partially due to population stratification. Genomic control (GC) approaches were proposed to detect and adjust for this confounder. To date, the performance of this strategy has not been extensively evaluated on real data. More than 10000 single-nucleotide polymorphisms (SNPs) were genotyped on subjects from four populations (including an Asian, an African-American and two Caucasian populations) using GeneChip(R) Mapping 10 K array. On these data, we tested the performance of two GC approaches in different scenarios including various numbers of GC markers and different degrees of population stratification. In the scenario of substantial population stratification, both GC approaches are sensitive using only 20-50 random SNPs, and the mixed subjects can be separated into homogeneous subgroups. In the scenario of moderate stratification, both GC approaches have poor sensitivities. However, the bias in association test can still be corrected even when no statistical significant population stratification is detected. We conducted extensive benchmark analyses on GC approaches using SNPs over the whole human genome. We found GC method can cluster subjects to homogeneous subgroups if there is a substantial difference in genetic background. The inflation factor, estimated by GC markers, can effectively adjust for the confounding effect of population stratification regardless of its extent. We also suggest that as low as 50 random SNPs with heterozygosity >40% should be sufficient as genomic controls.

Address: Wong, Wing H.; Sch Publ HlthDept Biostat, Harvard Univ, 655 Huntington Ave,Bldg 2,Room 441, Boston, MA, 02115, USA

ISSN: 1018-4813

[0 records marked]

(Save, Export, E-mail, Order, Print)

Create Citation Alert

Receive e-mail alerts on future citation to this record. (Requires registration.)

Additional Links

View in Web of Science

[Cited References](#)
[Citing Articles](#)
[Related Records](#)

View record in

[CAB Abstracts](#)
[Current Contents Connect](#)
[CC Connect Table of Contents](#)
[MEDLINE](#)
[Web of Science](#)

Explore articles from a variety of perspectives

These varied tools and options allow users to conduct research from many perspectives and in different contexts. For example, the *Affymetrix Genechip article's BIOSIS Previews[®]* record provides the life sciences view, including fields such as major concepts, concept codes, taxonomic data, and methods and equipment data.

For an agriculture-based angle, mutual subscribers to other *ISI Web of Knowledge* products can easily click on a link to the record in *CAB Abstracts[®]*; they can also view it in *MEDLINE[®]*,

Current Contents Connect[®], and *Web of Science*. Each product provides valuable descriptors, codes, identifiers and other data and search options that add to the user's understanding of this article's meaning and impact. And all records are displayed in one intuitive, easy-to-use interface.

Continued growth

BIOSIS continues to grow; over 168 journal titles have been added to its extensive coverage in the past year. Beth Ten Have, Senior Product Manager, Biosis, notes that the titles added represent a diverse group of subject and geographical areas:

"We are adding titles in the areas of agriculture, biotechnology, molecular biology, ecology and environmental management and conservation, and biomedicine. Over 20% of these titles are published in Asia, another 35% are from Western Europe, over 10% are from Eastern Europe, and approximately 9% of the titles are from South and Central America."

As Ten Have notes: "These upgrades and enhancements represent Thomson Scientific's continuing commitment to delivering the highest quality multidisciplinary and geographically varied content to its BIOSIS users."

see next page for figure 2

Figure 2: The fifteen millionth BIOSIS record, viewed in the *Web of Science* database

ISI Web of KnowledgeSM BIOSIS Previews

Web of Science[®]

Full Record

Record 1 of 1

Title: Detect and adjust for population stratification in population-based association study using genomic control markers: an application of Affymetrix Genechip (R) Human Mapping 10K array

Author(s): [Hao K](#), [Li C](#), [Rosenow C](#), [Wong WH](#)

Source: EUROPEAN JOURNAL OF HUMAN GENETICS 12 (12): 1001-1006 DEC 2004

Document Type: Article

Language: English

Cited References: [15](#) **Times Cited:** [1](#)

Abstract: Population-based association design is often compromised by false or nonreplicable findings, partially due to population stratification. Genomic control (GC) approaches were proposed to detect and adjust for this confounder. To date, the performance of this strategy has not been extensively evaluated on real data. More than 10000 single-nucleotide polymorphisms (SNPs) were genotyped on subjects from four populations (including an Asian, an African-American and two Caucasian populations) using GeneChip(R) Mapping 10 K array. On these data, we tested the performance of two GC approaches in different scenarios including various numbers of GC markers and different degrees of population stratification. In the scenario of substantial population stratification, both GC approaches are sensitive using only 20-50 random SNPs, and the mixed subjects can be separated into homogeneous subgroups. In the scenario of moderate stratification, both GC approaches have poor sensitivities. However, the bias in association test can still be corrected even when no statistical significant population stratification is detected. We conducted extensive benchmark analyses on GC approaches using SNPs over the whole human genome. We found GC method can cluster subjects to homogeneous subgroups if there is a substantial difference in genetic background. The inflation factor, estimated by GC markers, can effectively adjust for the confounding effect of population stratification regardless of its extent. We also suggest that as low as 50 random SNPs with heterozygosity >40% should be sufficient as genomic controls.

Author Keywords: population stratification; population-based study; association test; genomic control

KeyWords Plus: GENETIC ASSOCIATION; DIABETES-MELLITUS; DISEQUILIBRIUM; ADMIXTURE

Addresses: Wong WH (reprint author), Harvard Univ, Sch Publ Hlth, Dept Biostat, 655 Huntington Ave,Bldg 2,Room 441, Boston, MA 02115 USA
Harvard Univ, Sch Publ Hlth, Dept Biostat, Boston, MA 02115 USA
Dana Farber Canc Inst, Dept Biostat, Boston, MA 02115 USA
Affymetrix, Gen Collaborat, Santa Clara, CA USA
Harvard Univ, Dept Stat, Cambridge, MA 02138 USA

E-mail Addresses: wwong@hsph.harvard.edu

Publisher: NATURE PUBLISHING GROUP, MACMILLAN BUILDING, 4 CRINAN ST, LONDON N1 9XW, ENGLAND

Subject Category: BIOCHEMISTRY & MOLECULAR BIOLOGY; GENETICS & HEREDITY

IDS Number: 871VR

ISSN: 1018-4813

Record 1 of 1

[0 articles marked]

(Save, Export, E-mail, Order, Print)

Create Citation Alert

Receive e-mail alerts on future citations to this record.
(Requires registration.)

Additional Links

View record in

[BIOSIS Previews](#)
[CAB Abstracts](#)
[Current Contents Connect](#)
[CC Connect Table of Contents](#)
[Journal Citation Reports](#)
[MEDLINE](#)

[Acceptable Use Policy](#)
Copyright © 2005 The Thomson Corporation

Additional information

[Current Contents Connect](http://scientific.thomson.com/ccc/): <http://scientific.thomson.com/ccc/>

[ISI Web of Knowledge](http://isiwebofknowledge.com/): <http://isiwebofknowledge.com/>

[MEDLINE](http://scientific.thomson.com/products/medline/): <http://scientific.thomson.com/products/medline/>

[Web of Science](http://www.thomsonscientific.com/wos/): <http://www.thomsonscientific.com/wos/>