National Institute of General Medical Sciences

Outcomes Evaluation of the Large-Scale Collaborative Project Awards

Discovery Logic

THOMSON REUTERS

Dec 12, 2011

Prepared by:

Discovery Logic, a Thomson Reuters business
Rockville, Maryland 20850

Table of Contents

1.0 Executive Summary 3
2.0 Introduction 10
2.1 Glue Grant Program Goals 10
2.2 Evaluation Study Goals 11
2.3 Scope of Evaluation 11
3.0 Evaluation Design and Methodology 12
3.1 Evaluation Approach 12
3.2 Data Sources 12
4.0 Findings 17
4.1 Scientific Impact 17
4.2 Glue Grant Principal Investigator Funding 26
4.3 Collaboration and Shared Resources 30
5.0 Conclusions 45
6.0 Appendix: Supporting Information 48
6.1 Glue Grant Investigators 48
6.2 Collaboration Statistics for Glue Grant Investigators 51
6.3 Scientific Proximity Measurement 56

1.0 Executive Summary

The primary goal of the National Institute of General Sciences (NIGMS) Large-Scale Collaborative Project Awards (referred to in this document as the "Glue Grants" program) is to "enable the solution of major biomedical research problems and facilitate the next evolutionary stage of integrative biomedical science." Toward this end, the Glue Grants program funded five research consortia in an effort to create a more collaborative, interdisciplinary environment to advance research in key biomedical areas identified by the NIGMS research community.

This report outlines the results of a quantitative outcomes evaluation of the Glue Grants program. The five participating consortia are:

- Cellular Signaling (AFCS)
- Cell Migration (CMC)
- Functional Glycomics (CFG)
- Inflammation and the Host Response to Injury (IHRI)
- LIPID Metabolites and Pathway Strategy (LIPID MAPs)

Publications, citations, patents, clinical trials, funding, collaborations, and data bank depositions were examined from 2000-2011 and compared to both global benchmarks and topically similar NIH R01 awards. Following are the key findings:

(1) The Glue Grant program has had substantial overall scientific impact when compared to similar R01 awards, exceeding citation impact per publication over R01 comparison groups, with similar costs per citation within 6 years of the programs. This similarity in costs takes into account the large initial consortium costs compared with R01s.

- Publication velocity and productivity among Glue Grant recipients was qualitatively and quantitatively different from R01s funded during the same time period **(Figure E.1)**. In almost all cases, productivity reached that of matched R01 awards during the second 5-year award period. Publication citation impact in CMC, AFCS, and CFG was higher than in similar R01 groups. Patent awards for new molecules and resulting scientific uses were identified for two consortia (i.e., CMC, AFCS); the others had patent applications in process.
- IHRI met one of its stated goals for translational research by concluding two clinical trials during the Glue Grant period.
- Consortium publications exceeded their R01 comparison groups in observed over expected citations **(Figure E.2)**.

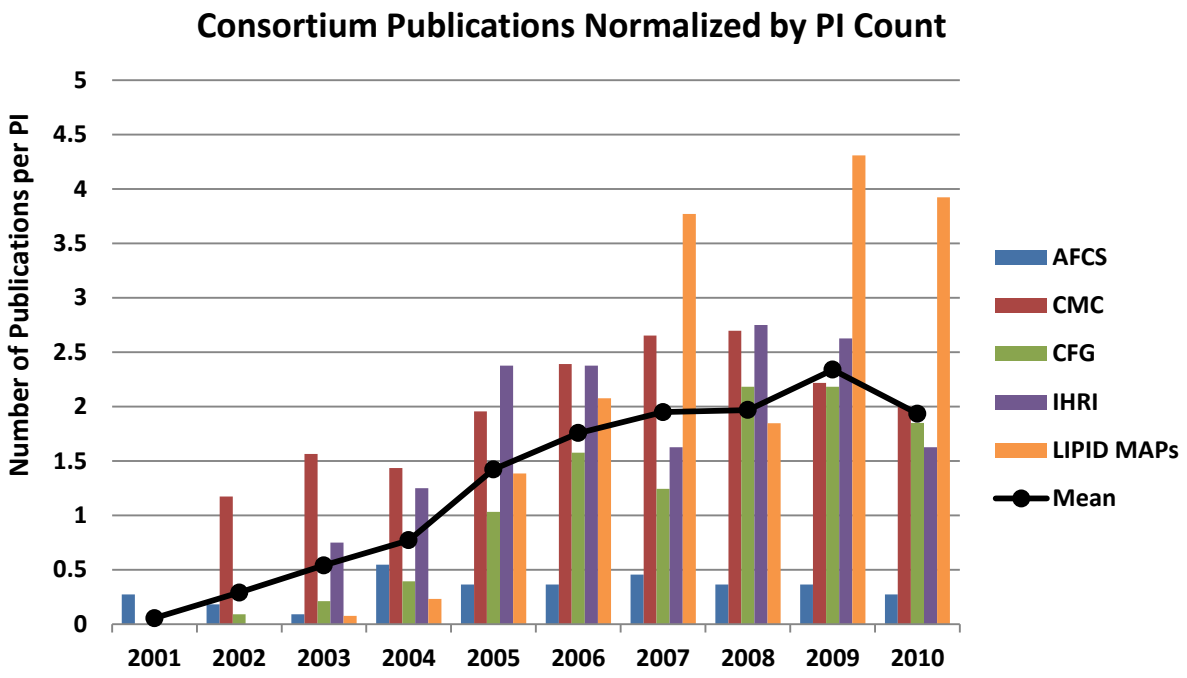


Figure E.1 Publications normalized by principal investigator for each consortium. Mean is defined as the annual mean publication count across the consortium, using the total consortium publication list. Although the number of investigators varied by year, only the unique total for each consortium was used to normalize the data. SOURCE: MEDLINE and Glue Grant progress report on self-reported publications.

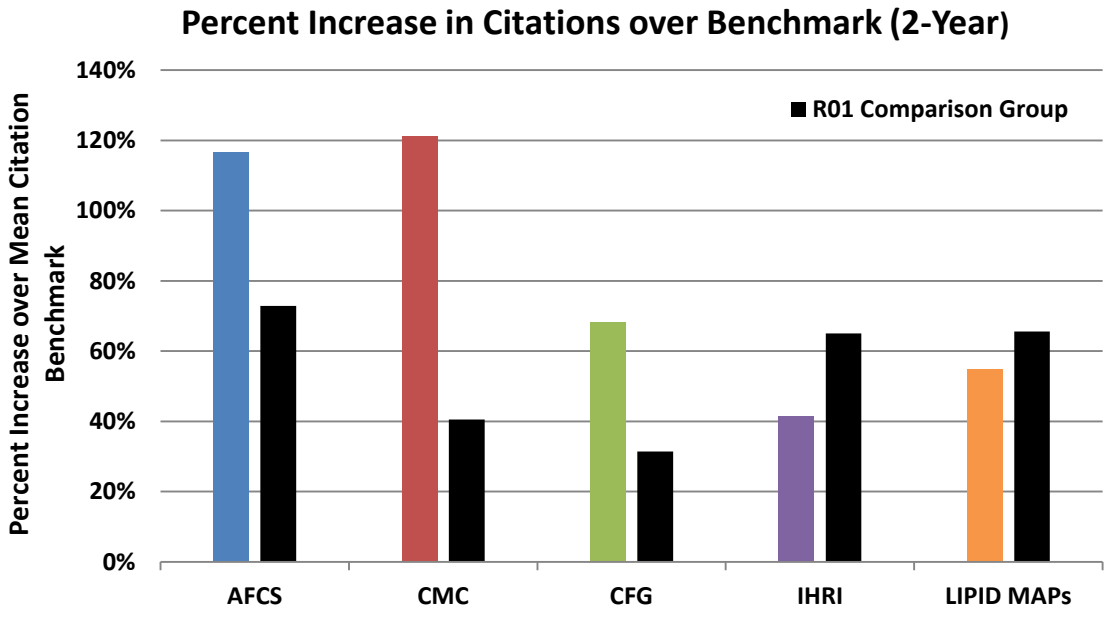


Figure E.2 The percent increase in citations observed over a 2-year period after publication compared to the benchmark expected citation rate for publications dated 2002-2008. SOURCE: MEDLINE and Thomson Reuters Web of Science

(2) The Glue Grants fostered increased scientific interaction among consortium investigators and increased interactions with the outside scientific community.

- Glue Grant investigators in total had more high-author count publications during the Glue Grant period than before and greater median values than the R01 comparison groups. Glue Grant publications in the LIPID MAPs, AFCS, and IHRI groups featured more authors per publication than publications by the same investigators outside the consortia **(Figure E.3)**.
- Twenty-four percent of publications acknowledging Glue Grant support featured no consortium PIs as authors. The majority of these were identified as core facility users.
- Shared data repository activity was a primary output for CMC and CFG groups, as observed by references from UniProt and the Protein Data Bank databases. These depositions were a greater part of consortium activity than for the R01, as determined by percentage of publications referenced.
- Glue Grant investigators from fields outside the pure biomedical sciences were observed to publish in more interdisciplinary publications during the Glue Grant period, as observed by overlap of MeSH terms with other Glue Grant investigators.
- Overall, scientific proximity between investigators significantly increased over time within most consortia (i.e., CMC, CFG, IHRI, LIPID MAPs), with the exception of investigators in AFCS.

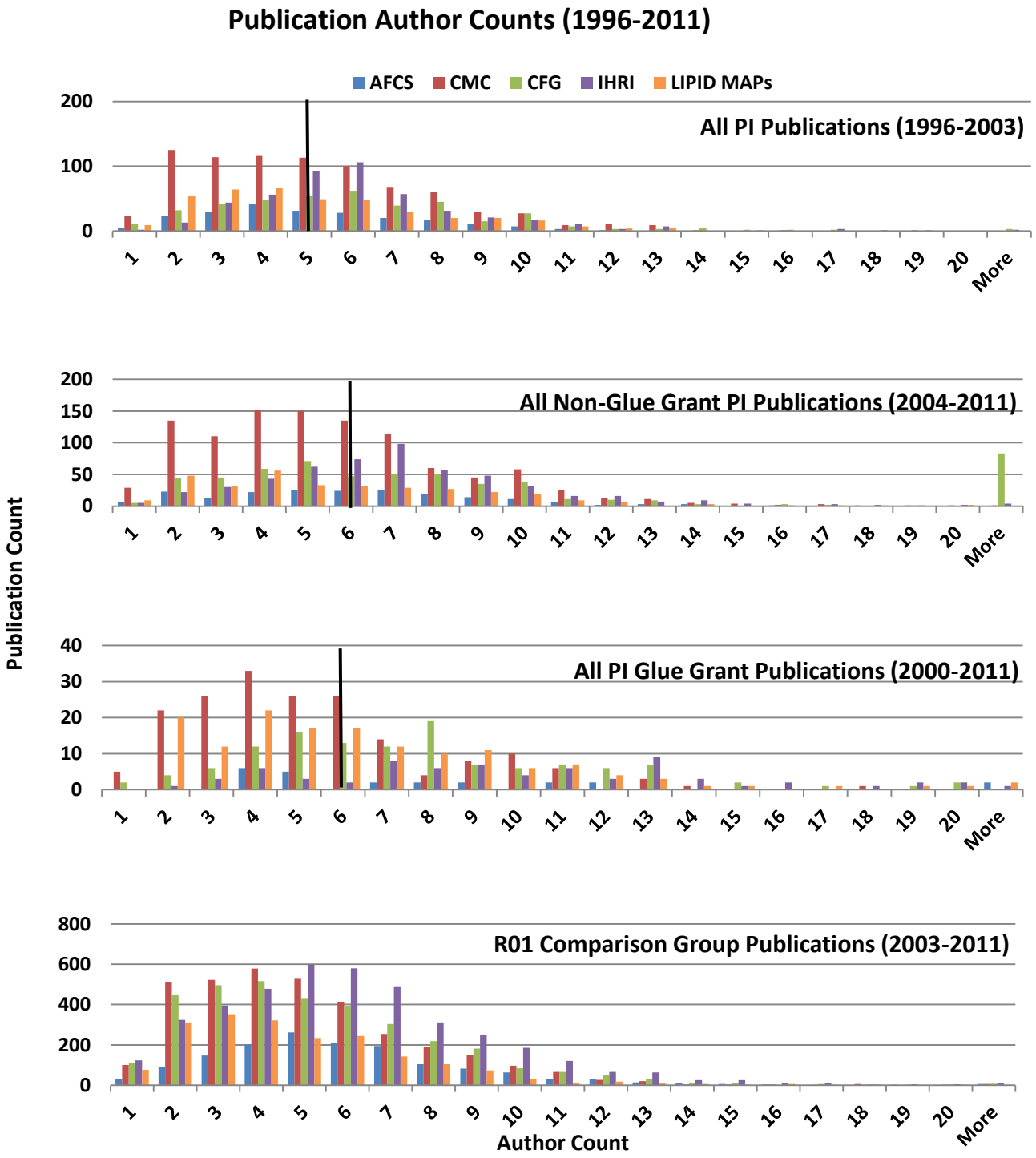


Figure E.3 **Number of authors for publications featuring at least one Glue Grant PI. (a) Investigator publications prior to the Glue Grant (1996-2003), (b) non-Glue Grant publications during the Glue Grant period (2004-2011), (c) Glue Grant publications (2000-2011), and (d) all publications produced by the R01 comparison groups. Black vertical lines represent the median value for all consortia in each group. SOURCE: MEDLINE and Thomson Reuters Web of Science**

(3) Investigators participating in the Glue Grant program were able to leverage their participation to win new NIH awards.

- The Glue Grants, on average, provided a significant fraction of funding for participating investigators. In four cases non-Glue Grant funding increased during the Glue Grant period (e.g., CMC, CFG, IHRI, LIPID MAPs). AFCS, however, saw little to no average change in overall funding outside the Glue Grants.
- Data suggest that investigators in all consortia except AFCS received increased funding during the Glue Grant period, particularly CFG. Additionally, co-funding outside the Glue Grants occurred between two investigators in CMC, which began in 2009 **(Figure E.4)**.
- The Glue Grant program fostered one new co-awarded NIH project grant between investigators in CMC, which began in 2009.

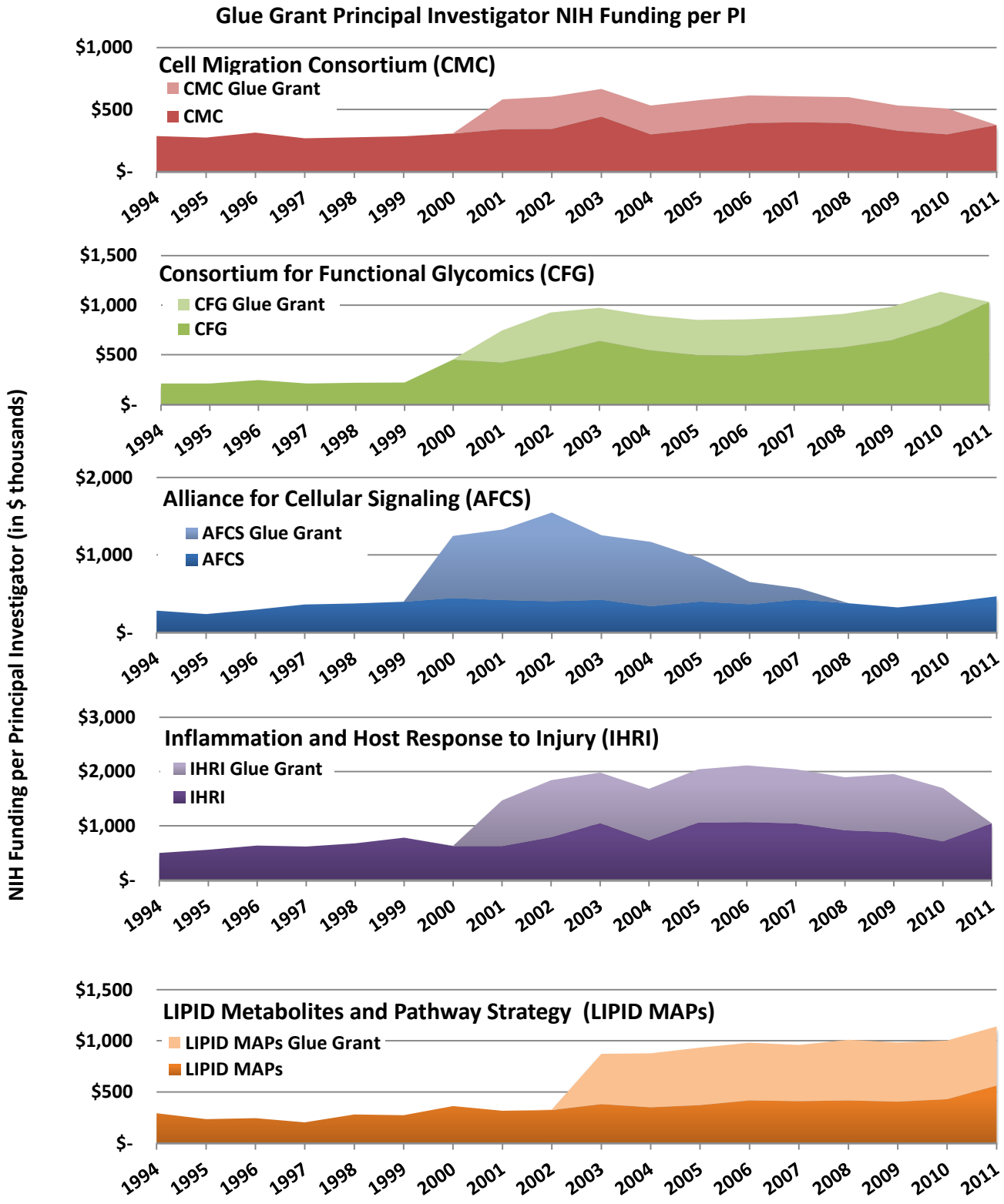


Figure E.4 Total NIH funding normalized by consortium PI. All funded award types that include one of the participating Glue Grant PIs as the award PI. Glue Grant funding is shown in the lighter shade overlaid on non-Glue Grant NIH funding in the darker shade. Normalization by PI used the same number of PIs throughout the period (11, AFCS; 23, CFG; 33, CMC; 8 IHRI; 13, LIPID MAPs). SOURCE: NIH CRISP and NIH RePORTer

(4) The data repositories created on the individual consortium websites were difficult to access or link to publications and individual contributors.

- Although shared data repositories were created by each consortium, they had insufficient accessibility to make evaluation of their impact feasible. No clear references to publications or investigators existed in an easily accessible data format for evaluation.

(5) Despite the focus of the consortia to establish collaborative research, it was unclear to what extent collaborations existed, except in the cases of co-authorship of publications acknowledging Glue Grant support.

- There was evidence that both a greater number of authors on average contributed to Glue Grant publications than R01 comparison groups and that greater collaboration across institutions existed.
- Information collected from grantees on collaborations with other principal investigators not directly receiving Glue Grant support would greatly facilitate future evaluations.

Overall, the Glue Grant program succeeded in fostering a collaborative environment focused on central biomedical problems. Increased scientific productivity, quality, and impact were observed in comparison to topically similar R01 comparison groups. Additionally, consortium activities resulted in a diverse set of targeted outputs, such as discovery of new protein information, filing of new patents, and completion of clinical trials. The level of success, however, varied by consortium.

The long-term impact on individual investigators is not clear at this time, but during the Glue Grant period, increased collaborative activity was observed, even for publications by Glue Grant investigators outside the consortium. Within consortium publications, a diverse collaborative environment was observed by institution, including institutions outside the United States.

Improved data collection from participants concerning collaboration, research products, and funding sources would enhance evaluation of similar programs in the future.

2.0 Introduction

2.1 Glue Grant Program Goals

The primary objective of the NIGMS Glue Grants program, initiated in 1999, is to enable the solution of major problems in biomedical research and to facilitate the next evolutionary stage of integrative biomedical science.[1] To achieve this objective, the Glue Grants program funded research consortia to create a more collaborative, interdisciplinary environment in which to advance research in key biomedical areas identified by the NIGMS research community over a 10-year timeframe.

The focus of individual consortia vary, ranging from standardizing practices in lipids research, to advance functional glycomics, to translating host response to injury research into clinical trial phases. Program goals include:

- **Cellular Signaling (AFCS)** to assess the complexity and dynamics of signal processing and to develop predictive models.
- **Cellular Migration (CMC)** to develop the tools and technology to help the cell migration research community elucidate the mechanisms underlying cell migration.
- **Functional Glycomics (CFG)** to define the biochemical mechanisms by which protein-carbohydrate interactions mediate cellular events.
- **Inflammation and the Host Response to Injury (IHRI)** research to improve the treatment of burn and trauma patients by applying high throughput biological discoveries to identify the dynamic relationships regulating the integration of the human immuno-inflammatory response.
- **LIPID Metabolites and Pathway Strategy (LIPID MAPs)** to apply an integrated approach to characterizing global changes in lipid metabolites, to establish the field of lipidomics, and ultimately, to aid in drug development for diseases in which lipids play critical roles.

While the scientific goals of each Glue Grant consortium are unique, the strategies used to achieve those goals are shared. These strategies include the implementation of core facilities to foster collaboration among researchers, the creation of multidisciplinary environments to more effectively solve specific research problems, and the shared dissemination of research advances in a systematic manner.

The NIGMS Glue Grant program distributed funding through a large-scale, collaborative U54 mechanism, in contrast to R01 awards targeted to individual researchers. This was done in part to foster better scientific collaborations among the participants, since the program was designed not only to advance research during the Glue Grant period, but to foster cross-disciplinary relationships and evolve an integrative scientific community that goes beyond the immediate scope of the Glue Grant.

[1] http://www.nigms.nih.gov/Research/FeaturedPrograms/Collaborative/GlueGrants/OutcomeAssessment/gluegrant_history.htm

2.2 Evaluation Study Goals

The purpose of this study is to provide objective, outcomes-based data on the effectiveness of the Glue Grant program when compared to traditional R01s. A central focus was the possible comparative advantage of the U54 mechanism over R01s in advancing major biomedical research areas. The evaluation is based on the overall program strategies of each Glue Grants consortium. As a result, the metrics used were designed to investigate the direct outcomes of Glue Grant funding, measure the indirect impact on research in the field, and observe the resulting scientific collaborations.

We measured relevant scientific outputs, such as publication productivity and citation impact, as well as the impact of participation on subsequent funding and collaboration. The goal was to determine whether the Glue Grants program fostered a collaborative, interdisciplinary environment for its investigators and whether there was increased interaction with similar researchers outside the consortium.

2.3 Scope of Evaluation

The evaluation compared the five Glue Grant consortia whose funding started in 2001 with a group of scientifically matched R01s active during the same period. The Glue Grant consortia did not all begin in the same year. Funding for AFCS was not renewed in the second 5-year period, but was provided with phase-down funding for the following 3 years. Principal investigators, lead institutions, and overall funding data are shown in **Table 2.1.** With the exception of LIPID MAPs (ending in 2013), the last funding year was 2011.

Table 2.1 **Glue Grant Consortium Summary**

Consortium Name	Lead PI	Lead Institution	Funding Period	Total Costs Through FY 2010 (millions)
Alliance for Cellular Signaling	Alfred Gilman	University of Texas Health Sciences Center, Dallas	2001-2007	$63.8
Cellular Migration	Alan Horwitz	University of Virginia	2001-2011	$79.9
Consortium for Functional Glycomics	James Paulson	The Scripps Research Institute	2001-2011	$80.2
Inflammation and Host Response to Injury	Ronald Tompkins	Massachusetts General Hospital	2001-2011	$78.4
LIPID MAPs	Edward Dennis	University of California, San Diego	2003-2013	$57.6

Section 3 describes evaluation design, indicators, data sources, and methodology. Section 4 presents the findings aligned with Glue Grant program goals and associated indicators, including the effects on subsequent investigator funding, publication productivity and impact, and collaborations and shared resources. In Section 5, overall study conclusions are presented. Section 6, the appendix, provides supporting information.

3.0 Evaluation Design and Methodology

3.1 Evaluation Approach

Because the specific scientific goals of the five Glue Grant consortia varied considerably, we chose to focus on assessing programmatic goals independent of research topics. Since a critical aspect of the Glue Grants program is to foster shared resources to facilitate breakthrough research, the evaluation was designed to assess the creation and use of shared resources. A comparison group of scientifically matched R01s also was created, in consultation with NIGMS program staff, to assess the impact of shared resources on scientific productivity and impact. The following questions guided our analysis:

(1) Funding Impact

- How effective are Glue Grant participants in obtaining subsequent *funding*?

(2) Scientific Impact

- How productive and influential are Glue Grant consortia with respect to *publications?*
- How productive are the Glue Grant participants with respect to *technology transfer* (i.e., invention disclosures applied for and patents received)?
- What evidence is there that Glue Grant research affects *clinical practices*?
- How effective are Glue Grant consortia in *establishing trends* for their field?

(3) Collaboration and Shared Resources

- Is the Glue Grant program increasing *collaboration* among consortium investigators?
- How productive are Glue Grant participants in contributing to *shared data repositories*? What is the *broader scientific engagement and impact* of these resources?
- Are Glue Grant consortium members conducting *interdisciplinary research*?

We used a number of indicators and data sources in this study, including publications and citations; previous and subsequent NIH grant awards; and data banks, patents, and clinical trials. In addition, we examined links between consortium publications and grants to other data sources, including patents and data banks. Patterns of co-authorship and use of subject area keywords were analyzed to quantify the level of collaboration and the breadth of expertise leveraged for research. Additionally, we linked publications and their citations to publicly shared data repositories, such as the Protein Data Bank, to measure public contribution and scientific impact.

3.2 Data Sources

3.2.1 Publications

Publication sets defined in this study included: (1) those directly linked to a consortium, (2) R01 comparison group publications, and (3) publications used to normalize impact and productivity to specific research areas. Publication data were collected from the National Library of Medicine's Medline and citation data were derived from Thomson Reuters' Web of Science Science Citation Index. We measured citations 2 years after publication date to normalize the length of time post-publication that

an article can accrue citations. Expected citations, or citations made to similar articles within the same journal and year, were used to benchmark performance.

For publications linked to Glue Grant activity, two partially overlapping publication sets were used **(Table 3.1)**. The first set, which we refer to as "self-reported," were submitted to NIGMS in the individual consortium progress reports.[2] The other set, which we refer to as "publicly available," were publications in the Medline database that directly acknowledged a Glue Grant number. "Both" refers to publications that appeared in both sets, while the combined, non-redundant set is represented by the "total" column. Note that cases in which the Glue Grants program is acknowledged by name rather than by grant number did not appear in the publicly available searches. Unless otherwise stated, the total publications were used to measure publication activity among the Glue Grant consortia.

Some publications appeared only in the self-reported list; typically because a grant number was not specifically acknowledged. A number of reasons account for this, but in most cases, it was mistakenly excluded or consortium investigators and/or researchers collaborated outside the consortium on consortium-relevant topics. Surprisingly, some publications appeared only on the publicly available list. Further investigation revealed that some of these publications provided grant number acknowledgement by core facility users or collaborators. Furthermore, many of the publications in both sets were not co-authored by any Glue Grant principal investigator.

Table 3.1 Glue Grant Publications and their Sources

Consortium	Self-Reported Only	Publicly Available Only	Both	Total
AFCS[2,3]	0	30	7	**37**
CMC	164	35	222	**421**
CFG	125	72	165	**362**
IHRI	10	73	42	**125**
LIPID MAPs	29	24	182	**235**
Total	**328**	**234**	**618**	**1180**

Publications co-authored by consortium investigators in the periods before and during the Glue Grant projects were identified using name-matching algorithms. Consortium investigators are defined as individuals whose names are listed on a grant application identified as part of the Glue Grants program and defined by the project number.[4] To determine the validity of these data, two measures of data quality were taken in comparison to a gold standard. The manner in which database matches are done typically results in identifying publications that are, to a high degree of confidence, co-authored by the

[2] "Self-reported" publication data for the AFCS are derived from the final progress report generated by NIGMS staff. This list is not intended to be exhaustive.
[3] AFCS did not have conventional publications as one of its major goals. Instead, the consortia focused on creating an online data repository as a means of communicating with the scientific community.
[4] For the full list of principal investigator names used in this study, see the Appendix in section 6.0.

individuals we desire, known as "precision." They are defined as the proportion of publications linked in the database that are also listed as publications by that individual meeting the gold standard measure. However, there are typically a significant number of publications that we fail to capture in our searches. The complement of this is known as "recall" and is measured as the proportion of publications listed in the gold standard that are also linked to that individual in the database. We assessed both precision and recall from a 5-10% sample of investigators by manual comparison to alternative sources, such as a curriculum vitae or a research website that lists publications. We found high precision and over 50% recall as expected **(Table 3.2)**, because the name-matching algorithms we used favor precision over recall. The level of error in recall is similar during the two periods of interest (within 10-15%), suggesting that outcome measurements before and during the Glue Grants program are subject to similar error and that it is appropriate to make comparisons between the time periods.

Table 3.2 **Precision and Recall Measurements for Glue Grant Investigator Publications**

Consortium	Overall	Overall	Before	Before Precision	During Recall	During Precision
AFCS PI	70.6%	90.8%	65.3%	92.5%	78.4%	88.9%
CMC PI	78.5%	95.2%	63.7%	93.7%	60.1%	97.1%
CFG PI	55.2%	96.0%	55.1%	97.4%	55.3%	95.0%
IHRI PI	54.0%	95.0%	48.9%	100.0%	59.8%	90.7%
LIPID PI	63.4%	94.2%	61.5%	95.5%	68.8%	93.1%
Mean	**64.3%**	**94.2%**	**58.9%**	**95.8%**	**64.5%**	**93.0%**

3.2.2 Funding

An initial set of the Glue Grant awards data was provided by NIGMS staff and included main PI names, project numbers, and funding amounts. These data were used to uniquely identify consortium investigators and to collect data on NIH funding awards other than Glue Grants for the fiscal years 1994-2011. A set of projects similar to each of the main topics of the Glue Grants was identified and used as a comparison group. Investigator and funding data for comparison group projects were identified using the same approach for each consortium.

3.2.3 Medical Subject Headings

Medical subject headings (MeSH) terms are used by the National Library of Medicine (NLM) to manually index publications in the Medline database, reflecting both subject areas covered in the publication (MeSH Descriptor) and further specification on the purpose and types of investigations (MeSH Qualifier). In this study, MeSH terms were used to better understand the topical focus of Glue Grant publications and those of the R01 comparison groups. They were also used to measure the professional relationships between Glue Grant investigators.

3.2.4 Patents

Patents citing Glue Grant support through grant number and consortium name were collected and evaluated using U.S. Patent and Trademark Office (USPTO) patent data from 2000. Patents that directly cited Glue Grant support were accessed through the same database.

3.2.5 Clinical Trials

ClinicalTrials.gov was used to identify trials with Glue Grant support by grant number and was confirmed by principal investigator through 2000. Current status and start and completion dates of clinical trials were also available through the public database, which was accessed on June 12th, 2011.

3.2.6 Shared Data Repositories

Deposition of protein structures and sequences from 2000 were obtained through the Protein Data Bank. Depositions were recorded by reference to Medline ID, which was then cross-referenced to the same ID of those publications citing Glue Grant support.

3.3 R01 Comparison Groups

A core component of this evaluation is determining the value of the Glue Grant program in comparison to individual investigator-initiated research grants (R01s) in advancing a large-scale area of research. Toward that end, a comparison group consisting of scientifically similar, NIGMS-funded R01 grants was generated for each consortium as a performance reference measure **(Table 3.3)**. Comparison projects were restricted to NIGMS-funded R01s with project activity during and after fiscal year 2000. R01 comparison groups were identified for the CFG and IHRI consortia using NIGMS program code classifications (PCC). Comparison R01 groups for the three remaining consortia were determined by keyword searches on project title, project abstract, and project terms, in consultation with NIGMS staff.

Table 3.3 **R01 Comparison Group Summary**

Consortium	Search Terms	Project Count	Pub Years	Mean Funding per Project per Year
AFCS	Keyword "macrophage" AND "signal transduction"	49	2001-2011	$211,251
CMC	Keyword "Cell Migration"	357	2001-2011	$175,209
CFG	P252 Project Code	127	2001-2011	$206,048
IHRI	P144 Project Code	266	2001-2011	$276,860
LIPID MAPs	Keyword ("lipid" AND "macrophage")	61	2001-2011	$199,721

The selected R01s were evaluated for similarity prior to inclusion in a comparison group. The R01 lists were manually reviewed by NIGMS staff to confirm relevance to the topic areas addressed by the consortia. Additionally, publication numbers were assessed to confirm that publication productivity was stable throughout the period of the Glue Grants **(Figure 3.1)**. Variations in publication numbers reflect not only the different numbers of projects and funding within each R01 comparison group, but also variations in the average numbers of publications per project **(Table 3.4)**. Other potential criteria, such as project start date, quality of individual researchers, and production history, were not considered in including or excluding projects from the R01 comparison group.[5]

[5] The observed lower mean publications per project for the CMC R01 comparison group has been attributed to an unusually high number of projects within the set that have no publications acknowledging grant support.

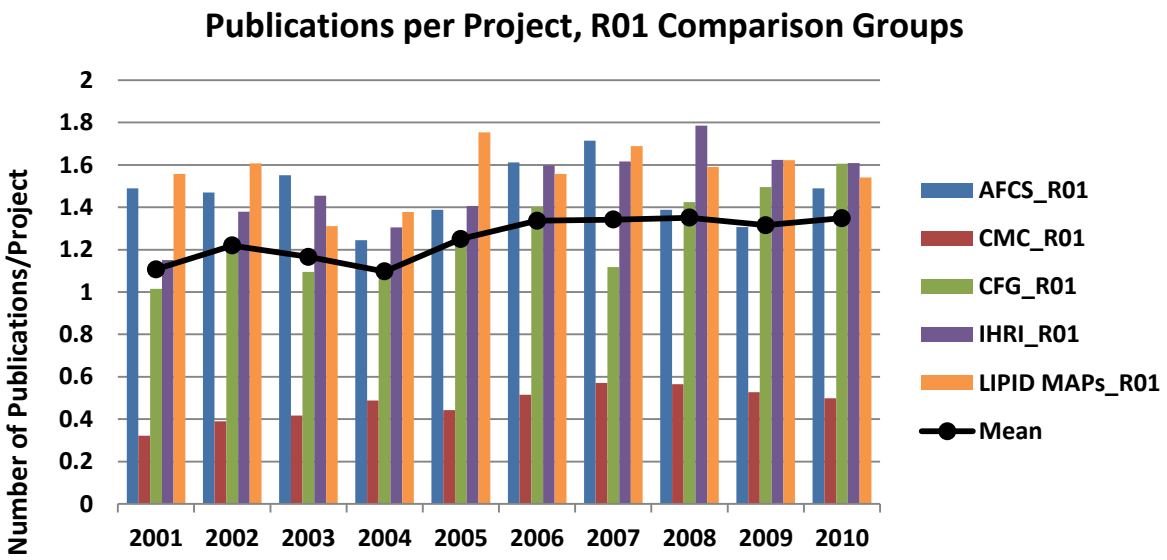


Figure 3.1 **Publication productivity by year for R01 comparison groups. Included are publications that acknowledged funding from NIGMS R01 projects that were active anytime between 2001 and 2011 and scientifically similar to the consortia, based on keyword search terms for the projects. The mean is the annual mean publications per project, measured across all R01 comparison groups. SOURCE: MEDLINE**

Table 3.4 **R01 Comparison Group Publication Counts**

R01 Comparison Group	Publication Count	Pubs/Project	Cost per Pub	Cost per Project
AFCS_R01	1494	30.5	$141,799	$4,324,870
CMC_R01	3482	9.8	$364,721	$3,574,266
CFG_R01	3370	26.5	$158,787	$4,207,856
IHRI_R01	4082	15.3	$183,196	$2,802,899
LIPID MAPs_R01	1956	32.1	$126,774	$4,069,445

4.0 Findings

4.1 Scientific Impact

Consortium goals varied, and consequently, so did their outputs. We assayed multiple outputs to measure productivity and impact. Productivity was measured in terms of publication volume, citations, quality of citations, scientific research topics, and trends. Other outputs measured included patent applications and awards and progress of research into clinical trials.

4.1.1 Publication Volume

As expected, Glue Grant publication volume varied substantially among consortia **(Figure 4.1)**. With the exception of CMC, there was a slow ramp-up in publication counts per year during the first 5-year grant period and stabilization during the second 5-year period.[6]

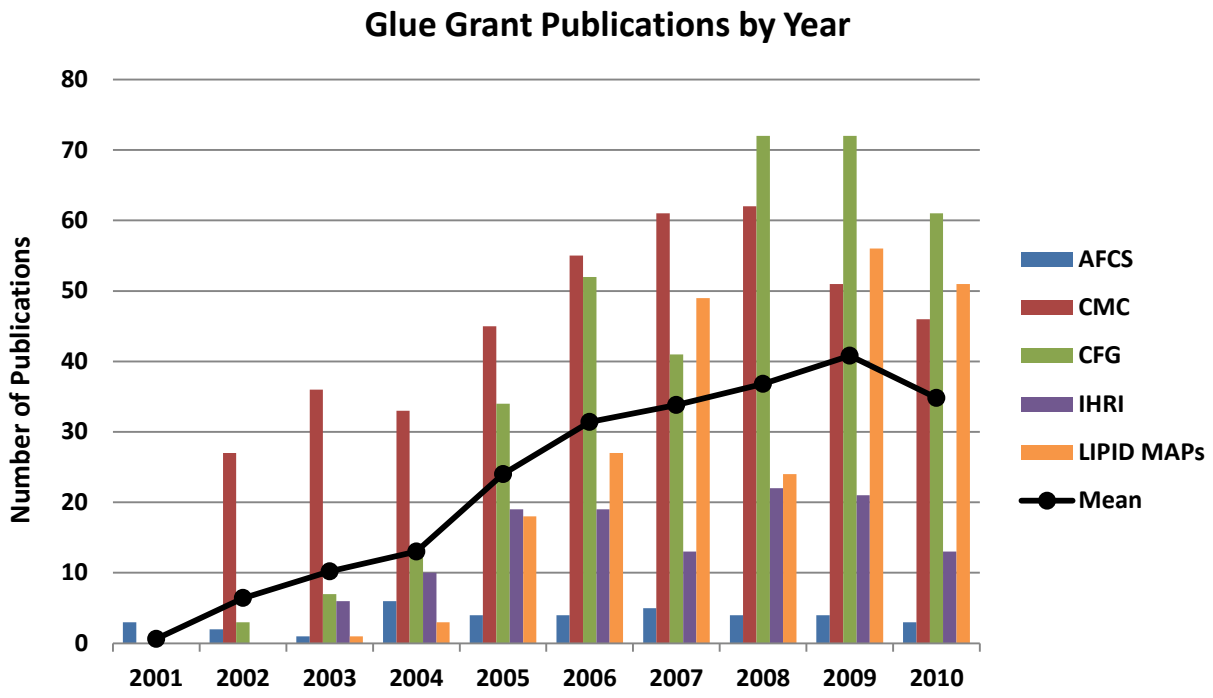


Figure 4.1 **Total publications by year shown for each consortium. The mean is the annual mean publication count across all consortia, using the total publication list. SOURCE: MEDLINE and Glue Grant Progress Reports.**

Variance in publication rates can be partially attributed to differences in the number of investigators and amount of funding across the consortia. Normalizing by number of participating PIs allows a comparison of relative publication productivity across all consortia **(Figure 4.2)**.The number of PIs was determined by the count of investigators appearing in at least one funded proposal in the Glue Grant project; all Glue Grant publications, regardless of authorship, were included. The mean is the annual

[6] It should be noted that the LIPID MAPS group first received funding in 2003, later than the other four consortia, while funding to AFCS was phased out beginning in 2005.

mean publication count per PI across all consortia, using the total publication list. Normalization by PI was performed by assuming that the same number of PIs were active throughout the Glue Grant period (11, AFCS; 23, CFG; 33, CMC; 8 IHRI; 13, LIPID MAPs).[7]

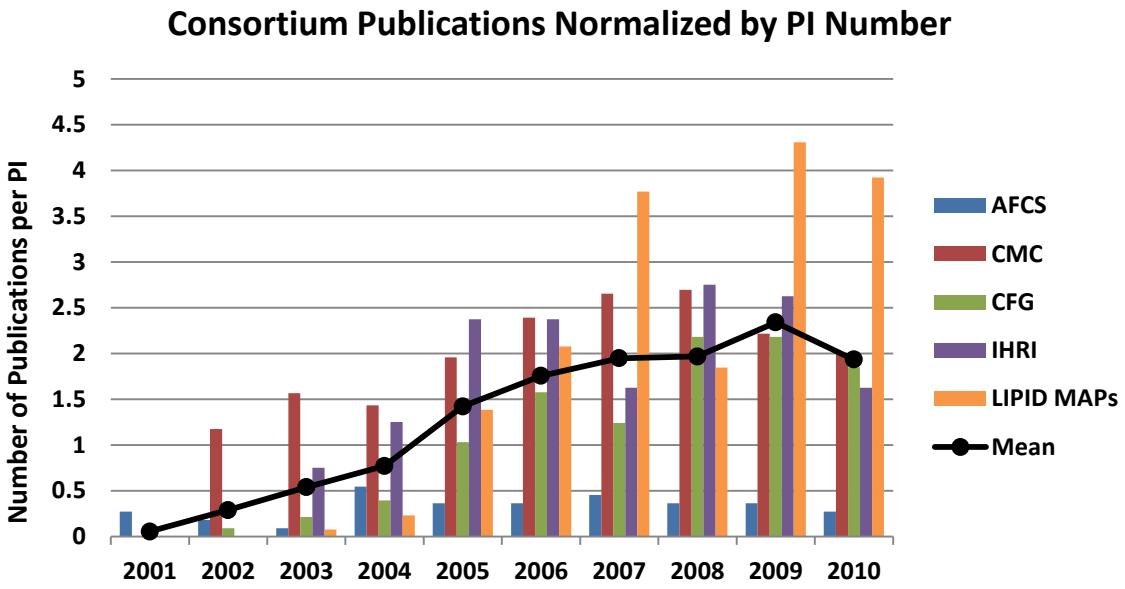
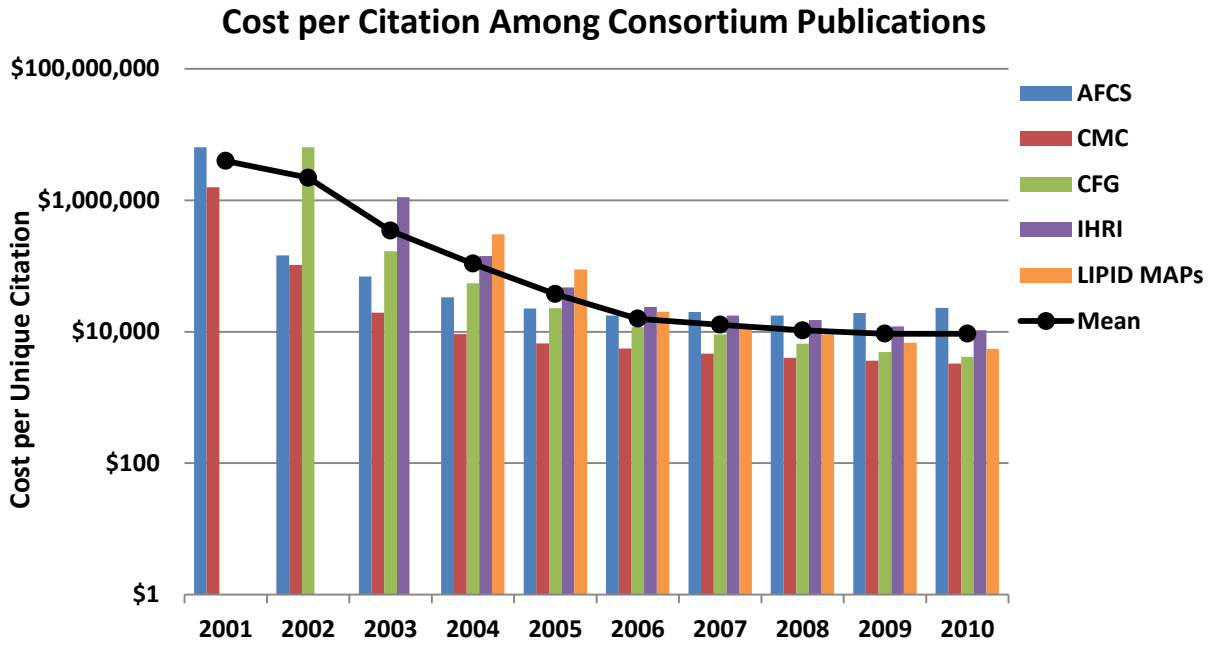


Figure 4.2 Publications normalized by principal investigator counts for each consortium. Although the number of investigators varied by year, only the unique total for each consortium was used to normalize the data. SOURCE: MEDLINE and Glue Grant Progress Reports.

4.1.2 Publication Impact

One of the most direct measures of publication impact is the number of times a publication is cited in reference sections of other publications. We therefore measured the unique citation counts per year for both consortia and R01 publications. Ultimately, impact needs to be measured against the program cost, and to properly compare the consortia impact with R01 impact, citations were normalized by cost per year **(Figure 4.3)**. In this case, the publications that cite Glue Grant publications are measured in the year of publication, normalized in each year by the program cost that year. This type of normalization was justified because program costs were relatively level throughout the Glue Grant period. As expected, the high initial costs of the consortia and low citation impact in early years resulted in significantly larger costs per citation during the initial years, compared with the R01 comparison groups. The costs, however, dropped by a rapid rate, and by 2007, reached levels similar to the R01 comparison groups. Once again, AFCS, due to its overall low publication counts, exhibited significantly higher costs per citation than the other consortia, especially by 2009.

[7] The assumption that the number of Glue Grant investigators was constant during the Glue Grant period applies to all normalizations by PI count in this report. For more detail on the investigators included, see Appendix section 6.1.



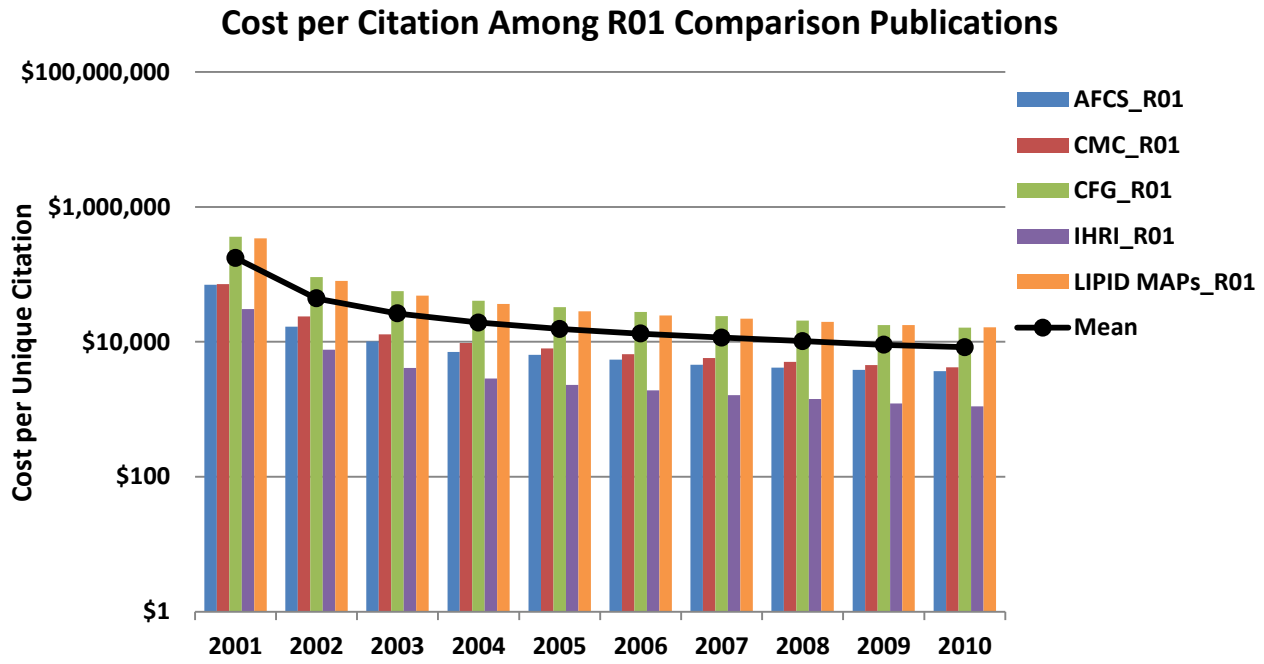


Figure 4.3 The cost per citation for Glue Grant publications (above) and R01 comparison groups (below), shown as a function of publication year. The costs are shown on a logarithmic scale for clarity. SOURCE: MEDLINE, Thomson Reuters Web of Science, and Glue Grant Progress Reports.

A modification of Thomson Reuters' Expected Citations[8] was used as a measure of citation performance over time, normalized by article type and journal.[9] The mean citation benchmark was calculated independently for each consortium and comparison group. Only publications that contained values for observed and expected citations were included.

The observed 2-year citation values represent the number of citations within 2 years of publication, while the observed 2-year not-self citations exclude any citations that have common authors among the cited publication. The observed citations of consortia and R01 publications exceeded the expected citations **(Table 4.1)**. The column, "2 Yr Observed/Benchmark Ratio" shows the ratio of the observed citations over their corresponding benchmarks. In relation to the R01 comparison group, three of the consortia had more citations and two had fewer citations **(Figure 4.4)**. Overall, the Glue Grants showed a larger difference from benchmark than the R01 comparison group.

Table 4.1 **Observed 2-Year Citations and their Citation Benchmarks for Glue Grant and Comparison Group Publications**

Consortium	2 Yr Observed Citations	2 Yr Observed Citations Not-Self	2 Yr Benchmark	2 Yr Benchmark Not-Self	2 Yr Observed/ Benchmark Ratio	Median Journal Impact Factor
Glue Grant Consortium						
AFCS	17.0	15.0	9.0	6.9	**2.2**	6.29
CMC	19.0	16.0	9.2	7.2	**2.2**	5.74
CFG	15.0	12.0	9.3	7.1	**1.7**	6.43
IHRI	13.0	10.0	8.9	7.1	**1.4**	3.32
LIPID MAPs	12.0	10.0	8.2	6.5	**1.5**	4.52
R01 Comparison						
AFCS_R01	11.3	9.3	7.0	5.4	**1.7**	4.57
CMC_R01	13.2	11.0	9.8	7.8	**1.4**	5.63
CFG_R01	10.2	7.6	7.8	5.8	**1.3**	6.29
IHRI_R01	8.9	6.7	5.4	4.0	**1.7**	3.68
LIPID MAPs_R01	12.3	9.6	7.8	5.8	**1.7**	4.76

[8] http://thomsonreuters.com/products_services/science/free/essays/expected_citation_rates/

[9] The citation benchmark assesses how a publication's citation rate performed against similar article types published in the same journal within 1 year of publication (from 6 months prior to 6 months after publication date). To further improve comparability across the data set, we normalized by limiting citation counts to those within 2 years from date of publication, from 2000-2008.

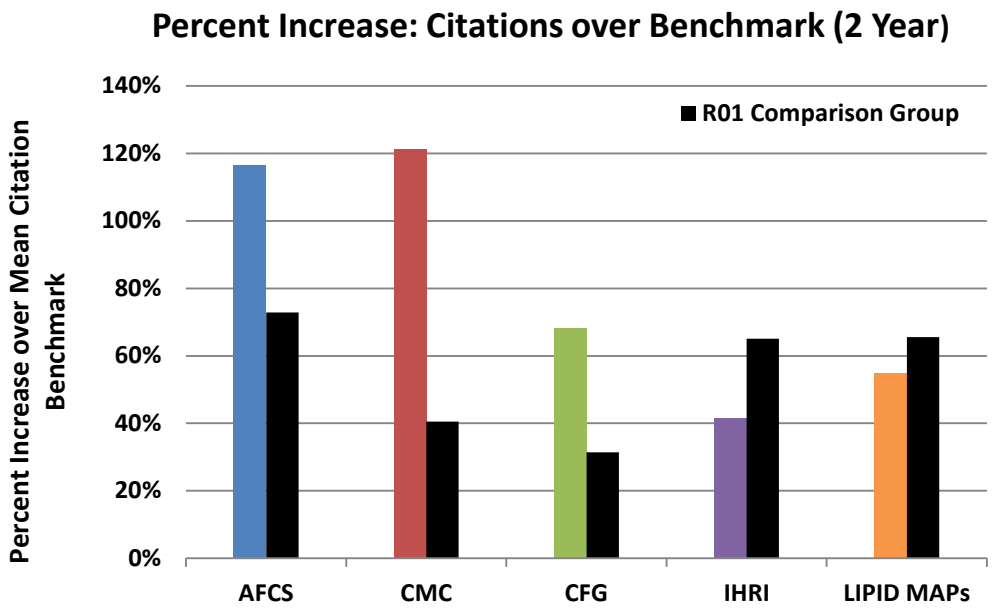


Figure 4.4 Percent increase of citations observed across a 2-year period after publication over its benchmark citation rate at 2 years after publication. Corresponding R01 comparison group results are shown in black for each consortium. SOURCE: Thomson Reuters Web of Science, 2002 to 2008

4.1.3 Publication Topic Areas

MeSH term occurrence among Glue Grant publications, shown in **Table 4.2,** provides insight into the scientific topic areas in which each consortium published, and into the relative focus of the topic compared to all research outputs in Medline. MeSH terms are divided into two broad categories: descriptors and qualifiers. Descriptors are scientific terms and topics that describe the content of publications' research, while qualifiers include methods and other terms to more precisely describe the type of research on which a publication is focused (e.g., metabolism, methods, chemistry, genetics). Qualifiers, additionally, are associated with a particular MeSH descriptor. For example, the descriptor "polysaccharides" can be associated with the qualifier "metabolism" or "chemistry," among others. The combination of the two types of terms provides more specific information on their use. In **Table 4.2**, frequencies were calculated on the paired use, rather than on each individually. Key MeSH descriptor/qualifier pairs found among consortium publications were identified by first filtering on those terms that appeared in at least 5% of a consortium's publications, then ranking those by the percentage of consortium publications to all of Medline (see last column in **Table 4.2**).

The majority of MeSH terms shared between consortium and R01 groups are related to scientific research topics, as opposed to instrumental and experimental methods. For example, one of the central goals of the LIPID MAPs consortium was to utilize mass spectrometry techniques in core facilities to advance research and standardization of lipids research. This is supported by the appearance of two mass spectrometry MeSH terms in the consortium's top 10. These same two terms, however, fail to appear in the top 10 list in the corresponding R01, suggesting that the consortium employed proportionally more resources to this technique in advancing the field. Similarly, the IHRI terms "Oligonucleotide Array Sequence Analysis" and "Gene Expression Profiling" are not in the R01 top 10.

Table 4.2 Top 10 MeSH Descriptors indexed to consortium publications.

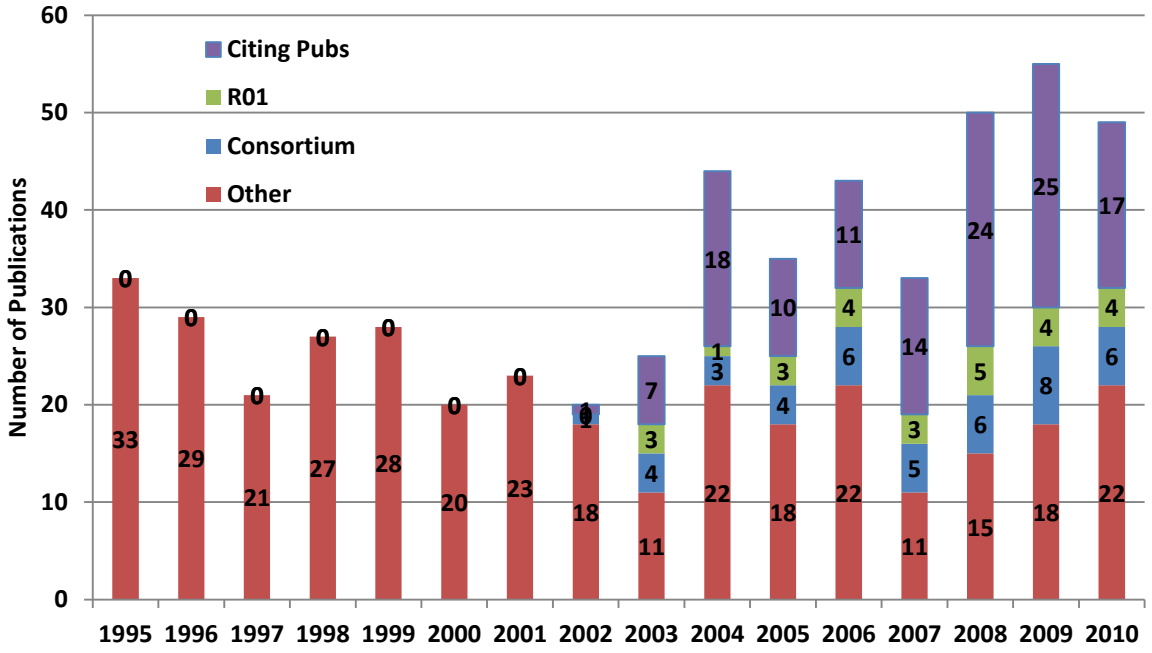
AFCS					
MeSH Descriptor	**MeSH Qualifier**	**Consortium Pub Count**	**MEDLINE Pubs**	**% of all Consortium Pubs**	**% MEDLINE**
UridineDiphosphate	pharmacology	3	79	8.11%	3.8%
Complement C5a	physiology	2	84	5.41%	2.4%
UridineDiphosphate	chemistry	2	87	5.41%	2.3%
Complement C5a	pharmacology	2	134	5.41%	1.5%
UridineDiphosphate	metabolism	2	140	5.41%	1.4%
Phospholipase C beta	metabolism	2	144	5.41%	1.4%
Cytological Techniques	instrumentation	2	153	5.41%	1.3%
Dinoprostone	immunology	2	204	5.41%	1.0%
Complement C5a	metabolism	2	206	5.41%	1.0%
Diglycerides	pharmacology	2	323	5.41%	0.6%
GTP-Binding Protein alpha Subunits, Gq-G11	metabolism	2	548	5.41%	0.4%
Cyclic AMP-Dependent Protein Kinases	genetics	2	749	5.41%	0.3%
GTP Phosphohydrolases	metabolism	4	1734	10.81%	0.2%
Phosphatidylinositols	metabolism	2	1237	5.41%	0.2%
GTPase-Activating Proteins	metabolism	2	1261	5.41%	0.2%
IHRI					
MeSH Descriptor	**MeSH Qualifier**	**Consortium Pub Count**	**MEDLINE Pubs**	**% of all Consortium Pubs**	**% MEDLINE**
Shock, Hemorrhagic	therapy	13	687	10.32%	1.9%
Burns	physiopathology	7	924	5.56%	0.8%
Multiple Organ Failure	etiology	9	1654	7.14%	0.5%
Resuscitation	methods	8	1521	6.35%	0.5%
Wounds and Injuries	complications	9	2240	7.14%	0.4%
Inflammation	genetics	10	3071	7.94%	0.3%
Wounds and Injuries	therapy	11	3957	8.73%	0.3%
Leukocytes	metabolism	7	3415	5.56%	0.2%
Proteome	analysis	7	4066	5.56%	0.2%
Proteomics	methods	12	10437	9.52%	0.1%
Oligonucleotide Array Sequence Analysis	methods	11	10638	8.73%	0.1%
Gene Expression Profiling	methods	11	10658	8.73%	0.1%
Inflammation	immunology	7	8328	5.56%	0.1%
LIPID MAPs					
MeSH Descriptor	**MeSH Qualifier**	**Consortium Pub Count**	**MEDLINE Pubs**	**% of all Consortium Pubs**	**% MEDLINE**
Sphingolipids	metabolism	12	754	5.08%	1.6%
Toll-Like Receptor 4	metabolism	12	1771	5.08%	0.7%
Phospholipids	metabolism	15	3768	6.36%	0.4%
Lipids	chemistry	22	6604	9.32%	0.3%
Spectrometry, Mass, Electrospray Ionization	methods	15	8414	6.36%	0.2%
Macrophages	metabolism	17	11976	7.20%	0.1%
Mass Spectrometry	methods	15	11527	6.36%	0.1%
Escherichia coli	metabolism	16	18594	6.78%	0.1%
Bacterial Proteins	metabolism	12	29134	5.08%	0.0%

CFG					
MeSH Descriptor	**MeSH Qualifier**	**Consortium Pub Count**	**MEDLINE Pubs**	**% of all Consortium Pubs**	**% MEDLINE**
Polysaccharides	metabolism	65	2679	18.01%	2.4%
Lectins	chemistry	32	1408	8.86%	2.3%
Polysaccharides	chemistry	62	4055	17.17%	1.5%
Lectins	metabolism	32	2161	8.86%	1.5%
Receptors, Virus	metabolism	20	1678	5.54%	1.2%
Oligosaccharides	metabolism	19	1979	5.26%	1.0%
Oligosaccharides	chemistry	20	3241	5.54%	0.6%
Receptors, Cell Surface	metabolism	19	9501	5.26%	0.2%

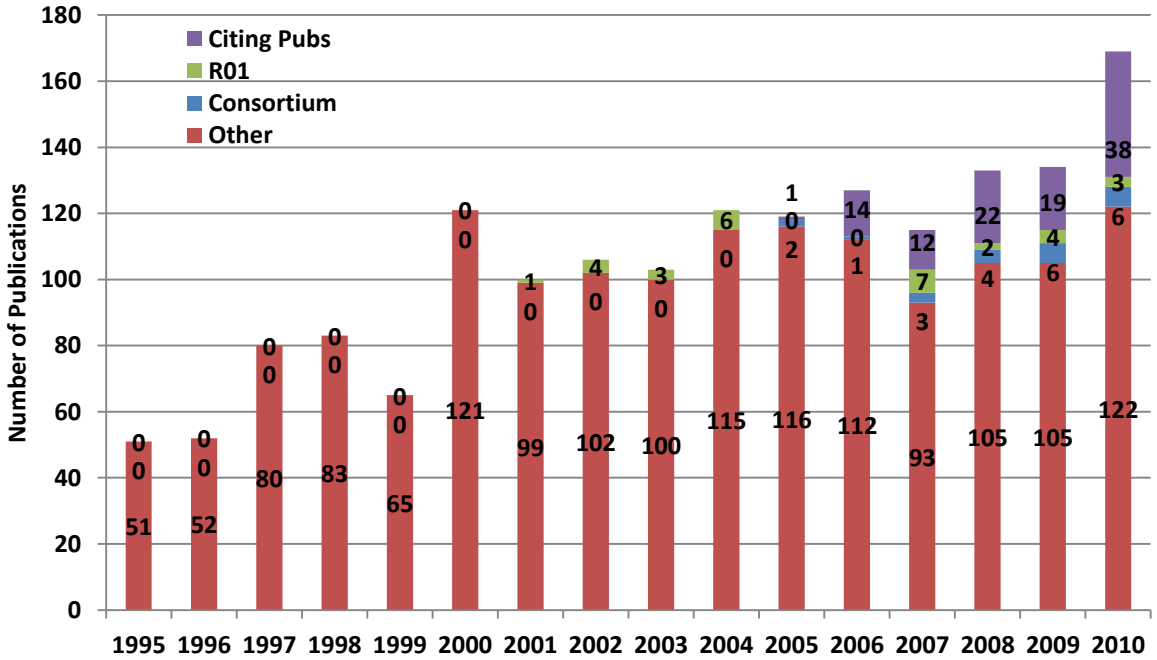
CMC					
MeSH Descriptor	**MeSH Qualifier**	**Consortium Pub Count**	**MEDLINE Pubs**	**% of all Consortium Pubs**	**% MEDLINE**
Talin	chemistry	29	94	6.81%	30.9%
Talin	metabolism	37	305	8.69%	12.1%
Pseudopodia	metabolism	29	911	6.81%	3.2%
Integrins	metabolism	40	3198	9.39%	1.3%
Microfilaments	metabolism	23	2131	5.40%	1.1%
Actins	chemistry	24	2355	5.63%	1.0%
Cell Movement	physiology	58	9752	13.62%	0.6%
Actins	metabolism	68	13460	15.96%	0.5%
Cytoskeleton	metabolism	28	6515	6.57%	0.4%
Cell Adhesion	physiology	28	6714	6.57%	0.4%
Green Fluorescent Proteins	metabolism	22	9803	5.16%	0.2%
Recombinant Fusion Proteins	metabolism	31	23597	7.28%	0.1%
Recombinant Fusion Proteins	genetics	22	18058	5.16%	0.1%
Peptides	chemistry	24	23123	5.63%	0.1%
Cell Membrane	metabolism	25	27380	5.87%	0.1%
Signal Transduction	physiology	32	43804	7.51%	0.1%

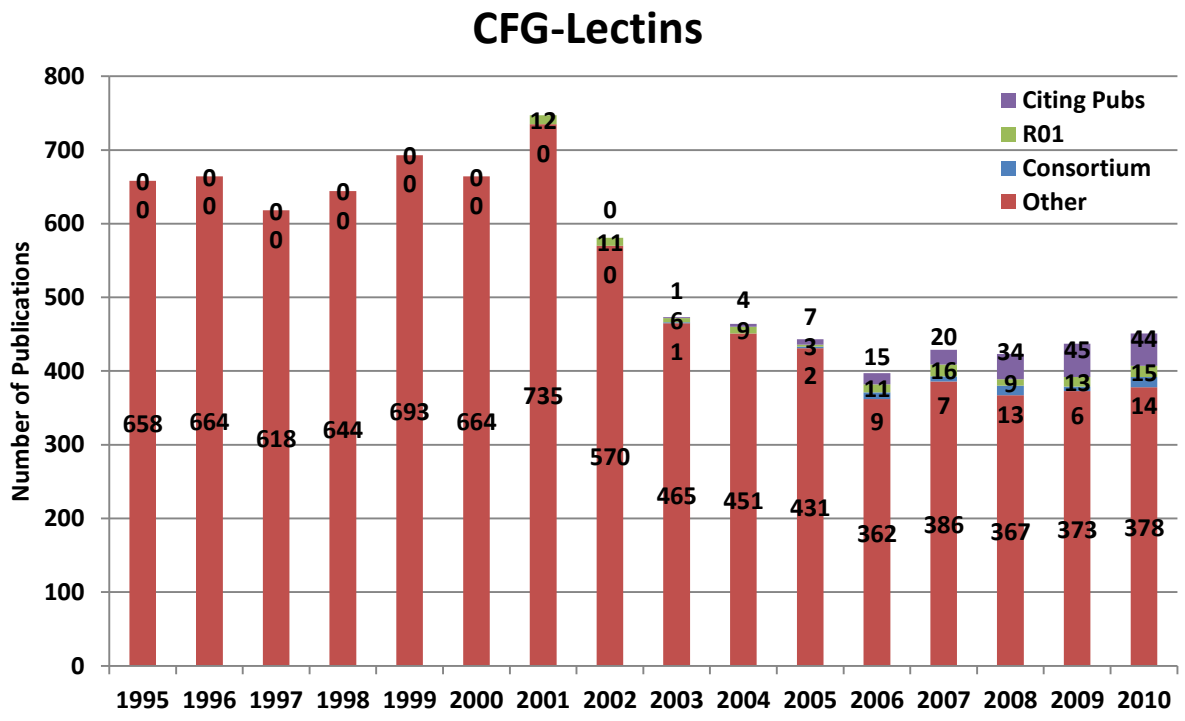
The MeSH term analysis reveals several key themes in a consortium's research publications. Measuring the occurrence of key descriptors over time provides additional information on the relative role the consortium played within its research field, as shown in **Figure 4.5**. Moreover, by measuring the MeSH occurrence among publications citing Glue Grant publications, we can examine the impact of the Glue Grant-funded research to the scientific community by MeSH term. The MeSH term counts are shown beginning in 1995 to demonstrate that the MeSH terms examined here existed prior to Glue Grant funding.

CMC-Talin



LIPID MAPs-Sphingolipids





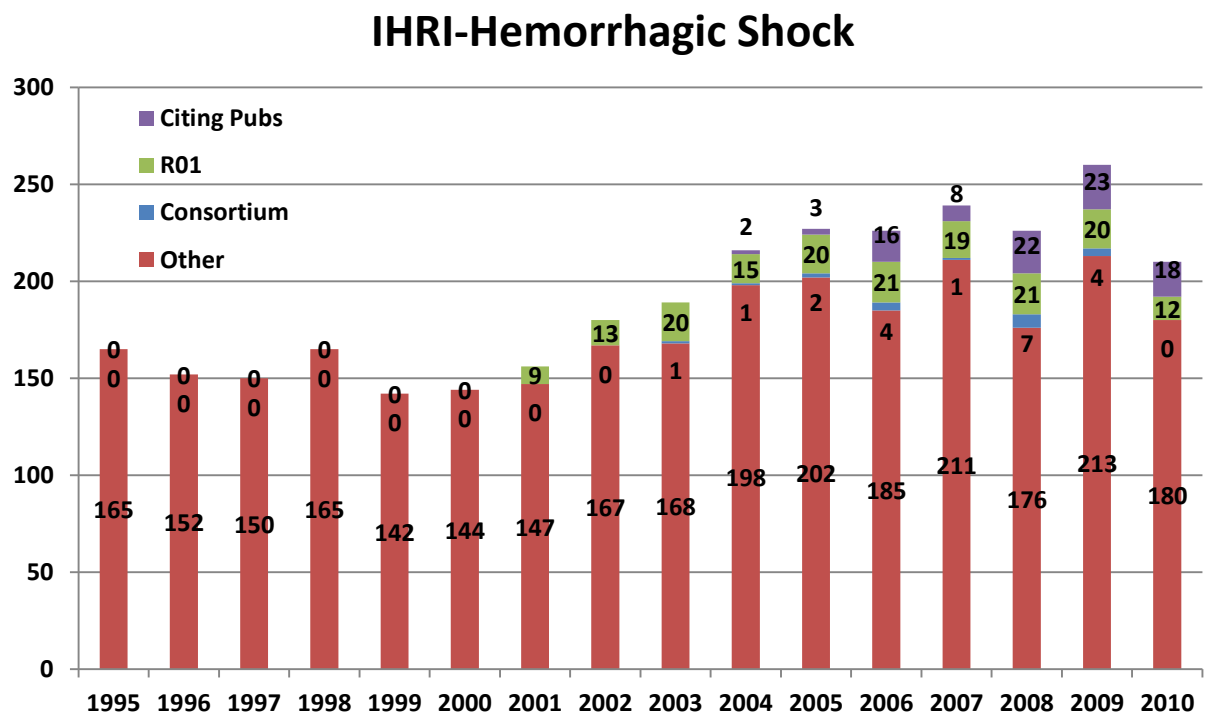


Figure 4.5 **Annual trends from 1995-2010 in usage of select MeSH descriptor terms, shown as total publications. Contributions are shown separately for publications citing Glue Grant support, publications citing R01 comparison group support, publications citing Glue Grant-supported publications (excluding publications in the other two groups), and all other publications in MEDLINE. MeSH Descriptors shown here are "Talin," "Sphingolipids," "Lectins," and "Shock, Hemorrhagic." AFCS MeSH Descriptors were left out due to low publication counts. SOURCE: MEDLINE**

4.1.4 Other Outputs

Research outputs other than publications were also considered, as they are an important component of the stated goals for some consortia. One of these outputs was Patents **(Table 4.3)**.[10] For the Inflammation and Host Response to Injury Consortium, a stated goal is translation of the research community's efforts through the clinical trial phases; thus clinical trials data (obtained from clinicaltrials.gov) provided a primary measure of output for this group **(Table 4.4)**.

Table 4.3 **Awarded Patents Citing Glue Grant Support. SOURCE: USPTO**

Consortium	Patents (USPTO)	Patent Applications (USPTO)	Patents Citing Glue Grant Patents
CMC	3	8	4
AFCS	8	10	11
CFG	0	3	0
IHRI	0	2	0
LIPID MAPS	0	0	0

Table 4.4 **Clinical Trials Citing Glue Grant Support (SOURCE: clinicaltrials.gov)**

	Principal Investigator	Start Date	Completion Date	Clinical Trials ID	Enrollment
IHRI	Ronald G. Tompkins	Nov 2003	Sept 2011	NCT00257231	610
IHRI	Ronald G. Tompkins	Apr 2004	Sept 2009	NCT00257244	280

4.2 Glue Grant Principal Investigator Funding

We examined total NIH funding, including all NIH grants, fellowships, and contracts, for all Glue Grant principal investigators.[11] The period from 1994 to 2002 was used as a pre-Glue Grant baseline. The years 2003-2011 were defined as the "Glue Grant period."

4.2.1 Total NIH Funding for Glue Grant Investigators

Grant funding awarded to Glue Grant investigators from 1994-2011 is shown in **Figure 4.6**. On average, funding outside direct Glue Grant support has remained relatively flat, although there are some substantial variations by consortium. The CFG group in particular has shown consistently significant gains in overall funding throughout the second 5-year Glue Grant period. Direct support from Glue Grants results in significant additions to the total NIH funding received by investigators in all cases, aside from CFG, which exhibited an upward trend in funding prior to receiving Glue Grant support. Again, aside from CFG, the total funding levels of the consortia investigators dropped off to roughly pre-Glue Grant levels at FY11.[12]

[10] Further detail on patents can be found in the Appendix (Section 6.0).
[11] See the Appendix (Section 6.0) for full list of principal investigators used in this analysis.
[12] Note that LIPID MAPs is operating until FY13.

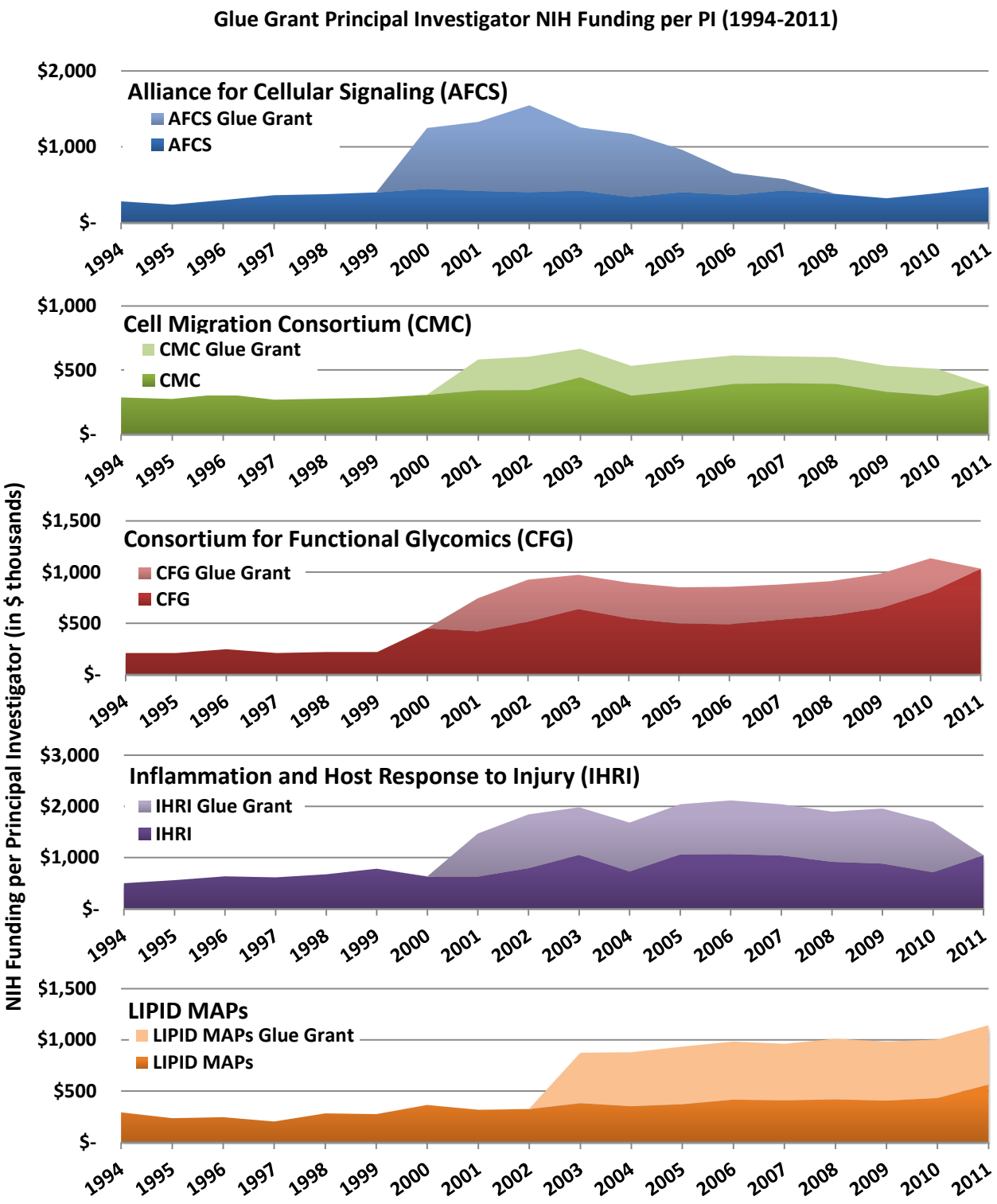
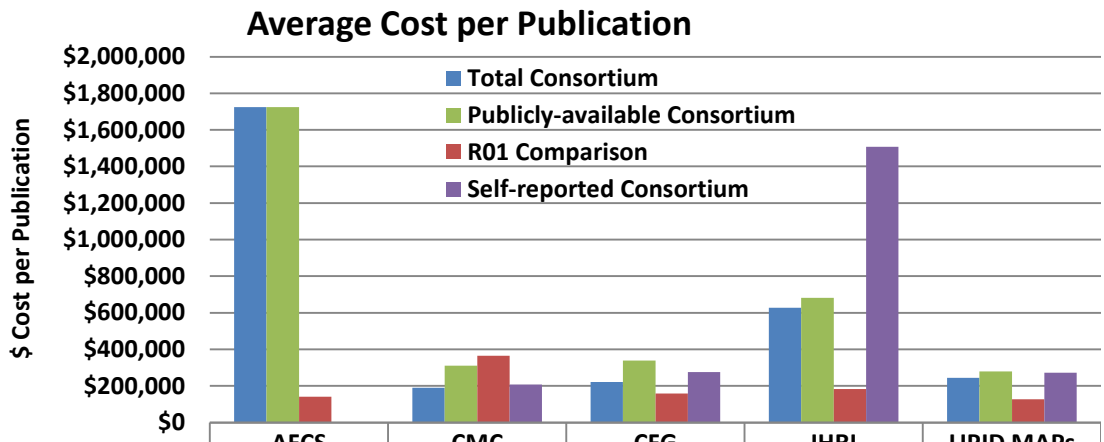


Figure 4.6 **Total NIH funding normalized by number of consortium PIs. All funded award types that include one of the participating Glue Grant PIs as the award PI. Glue Grant funding is shown overlaid (lighter shade) on all other NIH funding (darker shade). Normalization by PI assumed that each consortium had the same number of PIs throughout the period (11, AFCS; 23, CFG; 33, CMC; 8 IHRI; 13, LIPID MAPs). SOURCE: NIH RePORTer and NIH CRISP**

4.2.2 Cost per Publication

Consortium funding amounts were used to normalize publication counts. A commonly used metric for presenting this data is the cost per publication, for which the annual awarded funds are divided by the total publication output within the same time period. Publication and funding amounts for the same time period were considered, despite the common lag time between funds awarded and publication output resulting from those funds. This approximation is justified based on the relatively constant funding per year throughout the Glue Grant period for both the consortia and the R01 comparison groups.

As can be seen in **Figure 4.7**, three of the consortia (e.g., LIPID, CMC, and CFG) had costs comparable per publication to their R01 comparison groups (whose overall costs per project were significantly lower than the Glue Grants). Only the AFCS and IHRI consortia had average costs per publication exceeding their R01 comparison groups, which is consistent with their program goals in outputs other than research publications (among others, developing community tools and pushing research into clinical trials, respectively).



	AFCS	CMC	CFG	IHRI	LIPID MAPs
Total Consortium	**$1,724,324**	**$190,692**	**$221,547**	**$627,200**	**$245,106**
Publicly-available Consortium	**1,724,324**	**310,895**	**338,397**	**681,739**	**279,612**
R01 Comparison	**$141,799**	**$364,721**	**$158,787**	**$183,196**	**$126,774**
Self-reported Consortium	**$0**	**$206,995**	**$276,552**	**$1,507,692**	**$272,986**

Figure 4.7 **Cost per publication for each consortium and their R01 comparison groups. The total Glue Grant publications (blue), publically available publications (green), and the corresponding R01 comparison group publications (red) were used. Funding data was retrieved in publically available NIH project databases, while the publications from the consortia and their comparison groups were considered for the years 2001-2010.[13] SOURCE: MEDLINE, NIH RePORTer and NIH CRISP**

4.2.3 Other Funding Sources of Glue Grant Publications

Identifying the funding sources co-acknowledged in publications supported by Glue Grants provides some insight into the diversity of funding sources for consortium research. In **Table 4.5,** the top 5 non-

[13] The self-reported cost per publication value for AFCS is not shown here because the publication list is not intended to be exhaustive. See page 13 for further information.

NIH funding organizations, by number of publications acknowledging support, is shown for each consortium. For reference, **Table 4.6** shows the same data for the R01 comparison groups.

Although many of the large funders appear as expected on both Glue Grant and comparison group lists, including the National Science Foundation, American Heart Association, and the Wellcome Trust, there are a few notable differences. First, the CFG publications appear to be the most highly co-funded by a set of smaller funders, other than the Wellcome Trust, in comparison with the corresponding R01 group. Despite the low publication counts for AFCS, it is evident that the consortium received funding from a wide variety of institutions that are not traditional research funders.

Table 4.5 Non-NIH Funding Sources for Glue Grant Consortium Publications (2008-2011)

LIPID MAPs

Funding Organization	Pub Count (2008-2011)
National Science Foundation	9
American Heart Association	6
Leducq Fondation	5
Department of Energy	4
Natural Sciences and Engineering Research Council of Canada	3

CMC

Funding Organization	Pub Count (2008-2011)
Wellcome Trust	11
Israel Science Foundation	9
National Science Foundation	9
Cancer Research UK	8
American Heart Association	6

CFG

Funding Organization	Pub Count (2008-2011)
Wellcome Trust	11
Biotechnology and Biological Sciences Research Council	6
Skaggs Institute for Chemical Biology	4
CNRS	4
Singapore-MIT Alliance for Research and Technology	3

AFCS

Funding Organization	Pub count (2008-2011)
Microsoft Research	1
Yahoo! Research	1
Robert A. Welch Foundation	1
Department of Energy Computational Science Graduate Fellowship	1
Lockheed Martin Company	1
Ellison Medical Foundation	1
National Science Foundation	1
Fannie and John Hertz Foundation	1
Welch Foundation	1
Howard Hughes Medical Institute	1
Alfred and Mabel Gilman chair in molecular pharmacology	1
Laboratory Directed Research and Development program at Sandia National Laboratories	1

IHRI

Funding Organization	Pub Count (2008-2011)
Shriners Hospitals for Children	3
EPA	2
Entertainment Industry Foundation (EIF)	2
Battelle Memorial Institute for the DOE	2
NSF	2

Table 4.6 Non-NIH Funding Sources for R01 Comparison Group Publications (2008-2011)

LIPID MAPs

Funding Organization	Pub Count (2008-2011)
National Science Foundation	15
Perot Family Foundation	4
Department of Energy	4
Deutsche Forschungsgemeinschaft	4
Howard Hughes Medical Institute	3

CMC

Funding Organization	Pub Count (2008-2011)
American Heart Association	37
National Science Foundation	34
American Cancer Society	13
March of Dimes	6
Wellcome Trust	6

CFG

Funding Organization	Pub Count (2008-2011)
National Science Foundation	26
American Heart Association	14
Public Health Service	10
Wellcome Trust	8
Howard Hughes Medical Institute	7

AFCS

Funding Organization	Pub count (2008-2011)
American Heart Association	7
General Clinical Research Center	5
King Pharmaceuticals	5
Arthritis Foundation	4
Kaplan Cancer Center of New York University School of Medicine	3

IHRI

Funding Organization	Pub Count (2008-2011)
American Heart Association	64
Shriners Hospitals for Children	24
Shriners of North America	15
National Science Foundation	13
Department of Defense	9
American College of Surgeons	9

4.3 Collaboration and Shared Resources

4.3.1 Extent of Researcher Collaboration

One of the central goals of the Glue Grants is to foster a more collaborative and interdisciplinary research environment as a more effective means of solving key biological problems. Collaboration was measured using PI co-authorship on publications and co-funding on grants. Co-authorship can occur between PIs at the same or different institutions; the distinction between institutions was not considered for this analysis. Collaboration through grant co-funding occurs when two collaborators apply and are awarded an NIH grant as multi-PIs.[14]

One measure of collaboration can be made by identifying co-funded NIH research awards. Here we examined this type of funding collaboration for Glue Grant investigators. Because this model was adopted by the NIH in 2007, we expect to find these occurrences only from 2007-2011. **Table 4.7** shows total projects linked to Glue Grant PIs, overall multiple-PI project counts (defined as projects with more than one PI) among at least one Glue Grant investigator, and multiple-PI project counts among more than one Glue Grant investigator. The one observed case in which two Glue Grant PIs were co-awarded an NIH grant during the Glue Grant period began in 2009 and is scheduled to be completed in 2014.

[14] The NIH began implementing the multiple PI model in its research awards in FY07. See http://grants.nih.gov/grants/multi_PI/ for more information.

Table 4.7 **NIH Co-funded projects. SOURCE: NIH RePORTer and NIH CRISP**

Glue Grant Investigator NIH Projects (1994-2011)			
	Total Projects	Projects >1 PI	Projects by Glue Grant Investigators
Projects associated with Glue Grant Investigators	518	21	1[15]

We examined collaborative research activity using co-authorship metrics **(Figure 4.8)**. We found that the majority of publications acknowledging consortium support have one Glue Grant author (54% overall). A substantial proportion of publications have two Glue Grant PI authors (18% overall). A small number of publications listed five or more PI authors; upon manual review, these tended to be publications announcing new classification systems (e.g., LIPID MAPs), public data repositories, or other consortium-wide research.

Twenty-four percent of Glue Grant publications do not feature a Glue Grant PI as an author. In fact, CFG has more publications with no PIs than publications with one. This suggests that a significant proportion of publications citing Glue Grant support were not members of the consortia. This was confirmed by manual inspection of funding acknowledgements text in a sample of 10 publications for each consortium, to be a result of core facility use, materials or data transfer, or other forms of scientific interaction.

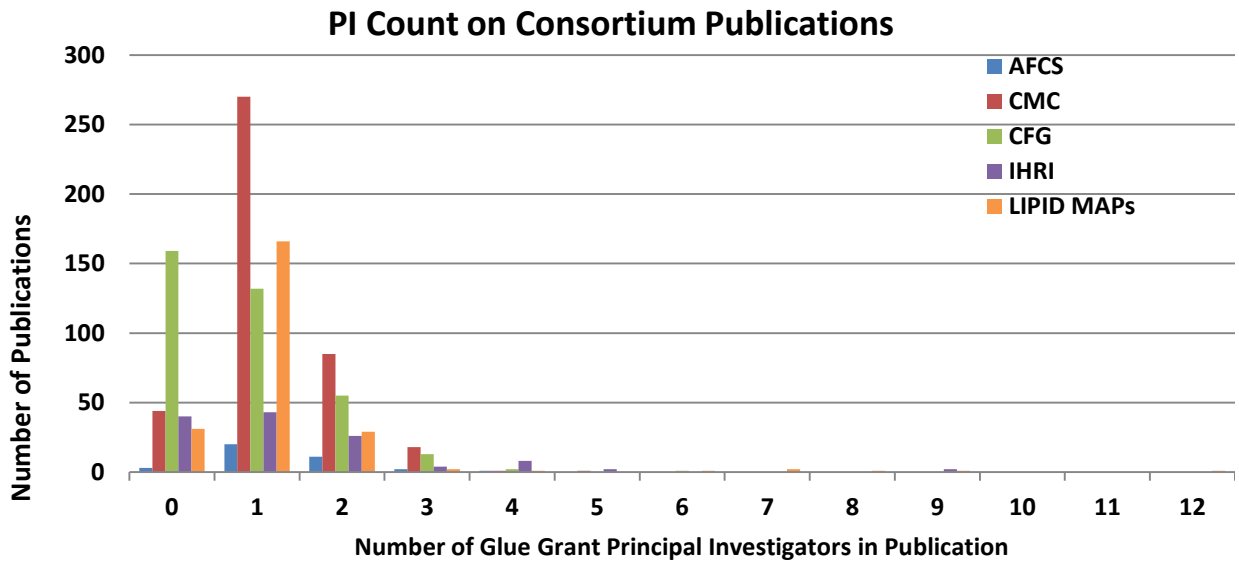


Figure 4.8 **Distribution of the number of principal investigators for Glue Grant publications, 2001-2011. Publications include both publically available and self-reported Glue Grant publications for which a Medline ID existed. SOURCE: MEDLINE and Glue Grant Progress Reports.**

Next, we investigated co-authorship among all researchers within a broader set of publications featuring at least one Glue Grant PI from 1996 to 2011. We compared investigator publications prior to the Glue Grant period (1996-2003), during the Glue Grant period (2004-2011), all Glue Grant publications (2000-

[15] The project co-funded between investigators in the Cellular Migration Consortium began in 2009.

2011), and all publications from the corresponding R01 comparison groups (2001-2011). Overall means across all consortia were calculated for each set.

The distributions in **Figure 4.9** and medians in **Table 4.8** show that there is an overall shift to more highly collaborative publications from the prior period to the Glue Grant period.

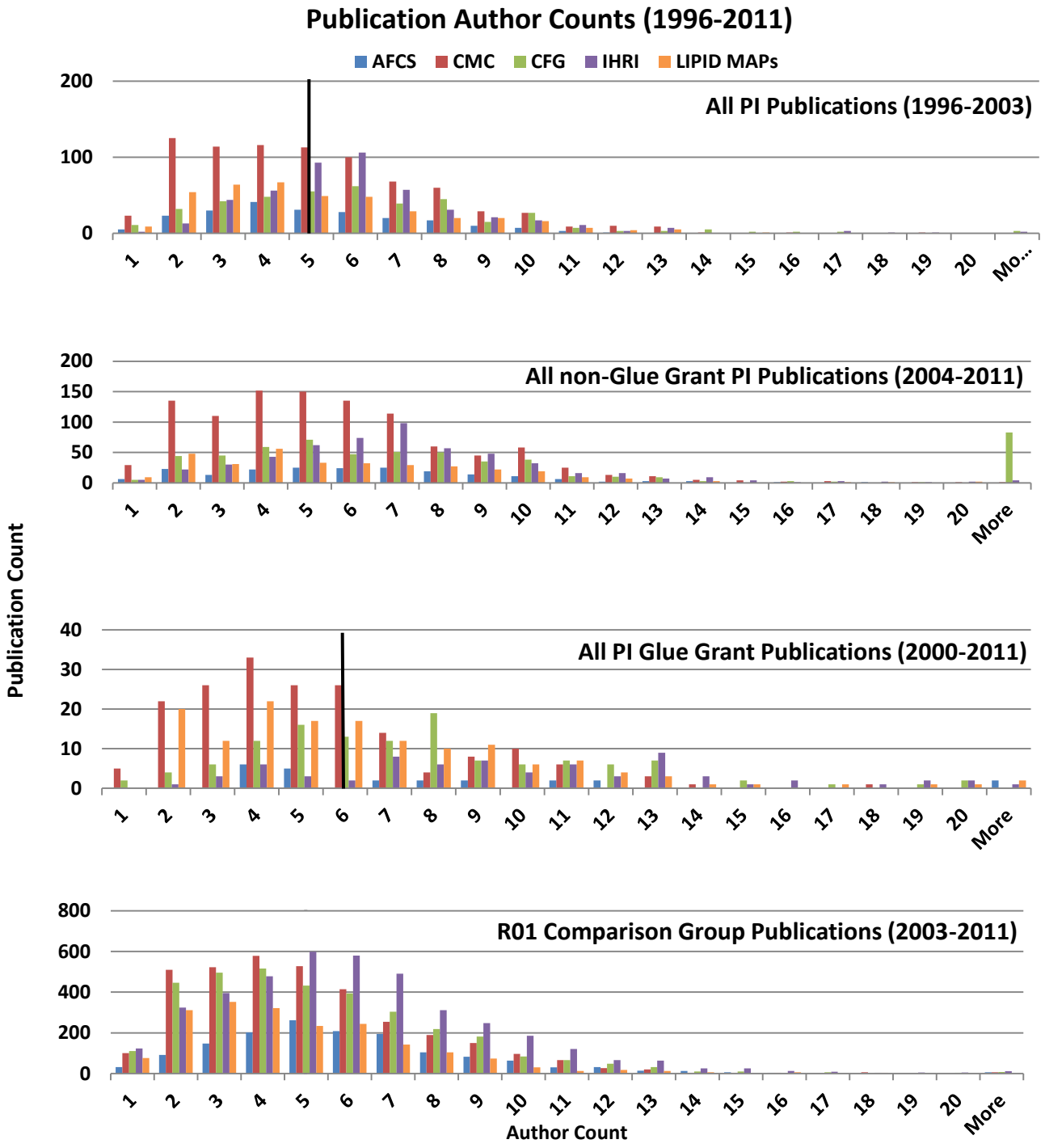


Figure 4.9 Number of authors for publications featuring at least one Glue Grant PI. Shown in four sections, (a) Investigator publications prior to the Glue Grant (1996-2003), (b) non-Glue Grant publications during the Glue Grant period (2004-2011), (c) Glue Grant publications (2000-2011), and (d) all publications produced by the R01 comparison groups. Black vertical lines represent the median value for all consortia in each group. SOURCE: MEDLINE

In making individual comparisons of median author count between specific groups **(Table 4.8)**, it becomes evident that overall, the Glue Grants had a higher median author count than their R01 comparisons groups. The most notable differences were in the IHRI, CFG, and LIPID MAPs consortia.

Furthermore, the median author count increased for PI publications from before to during the Glue Grant period (from 5 to 6), above the R01 groups, suggesting that Glue Grant PIs had an overall increase in author count compared with similar researchers during the Glue Grant period. Variations among individual consortium are observed, with IHRI investigators showing the most dramatic overall increase.

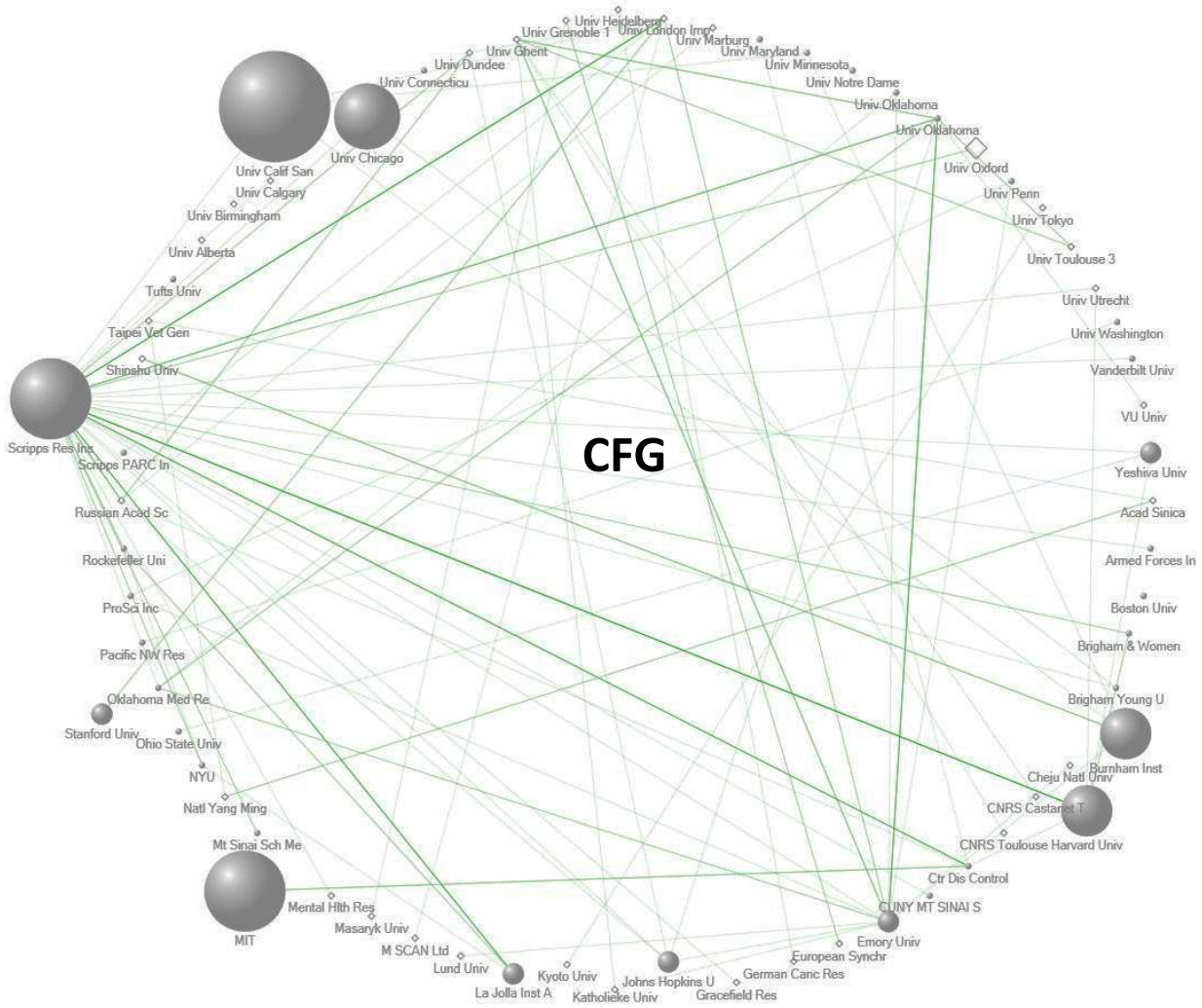
Table 4.8 Comparison of median author count for R01 publications vs. Glue Grant publications, Glue Grant publications in comparison to publications by the same investigators in the period prior to the Glue Grant, and Glue Grants versus other publications by the same investigators during the period of the Glue Grant. T-tests for R01 comparisons were performed using a 2-sample unpaired t-test, while that for Glue Grant vs. prior publications and Glue Grant vs. during publications performed using a 2-sample paired t-test.

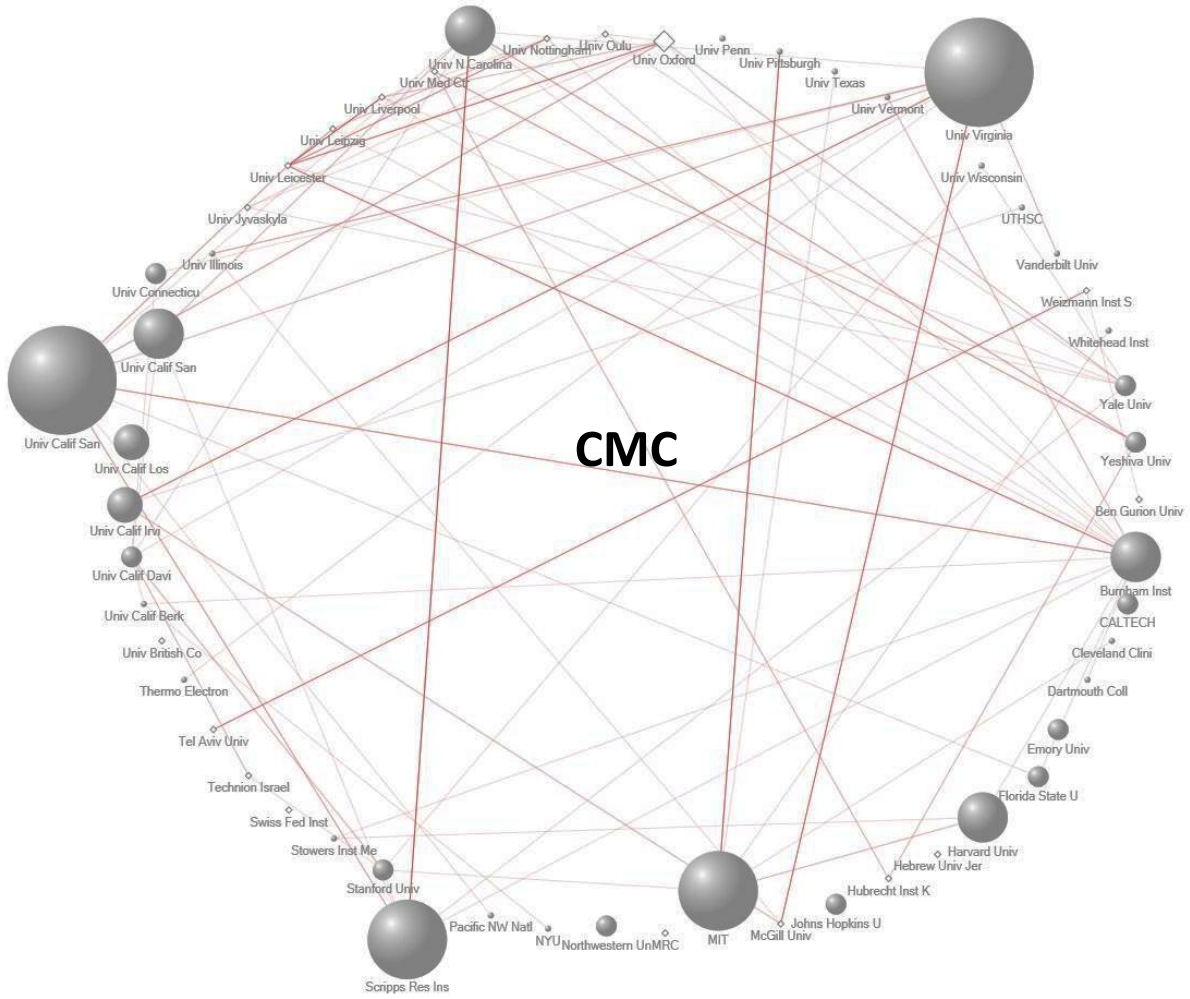
Publication Sets		Median	Median	P-value
R01 vs. Glue Grant		**R01**	**Glue**	**p-value**
	AFCS	6.0	7.0	0.0002
	CMC	5.0	5.0	0.19
	CFG	5.0	7.0	<0.0001
	IHRI	6.0	9.0	<0.0001
	LIPID MAPs	4.0	6.0	<0.0001
	Overall	***5.0***	***6.0***	***<0.0001***
PI Prior vs. Glue Grant		**Prior**	**Glue**	
	AFCS	5.0	7.0	<0.0001
	CMC	5.0	5.0	0.17
	CFG	6.0	7.0	0.0003
	IHRI	6.0	9.0	<0.0001
	LIPID MAPs	5.0	6.0	<0.0001
	Overall	***5.0***	***6.0***	***<0.0001***
PI During vs. Glue Grant		**During**	**Glue**	
	AFCS	6.0	7.0	0.002
	CMC	5.0	5.0	0.12
	CFG	7.0	7.0	0.0005
	IHRI	7.0	9.0	<0.0001
	LIPID MAPs	5.0	6.0	0.008
	Overall	***6.0***	***6.0***	***0.16***
PI Prior vs. PI During		**Prior**	**During**	
	AFCS	5.0	6.0	0.0001
	CMC	5.0	5.0	0.0001
	CFG	6.0	7.0	<0.0001
	IHRI	6.0	7.0	<0.0001
	LIPID MAPs	5.0	5.0	0.004
	Overall	***5.0***	***6.0***	**<0.0001**

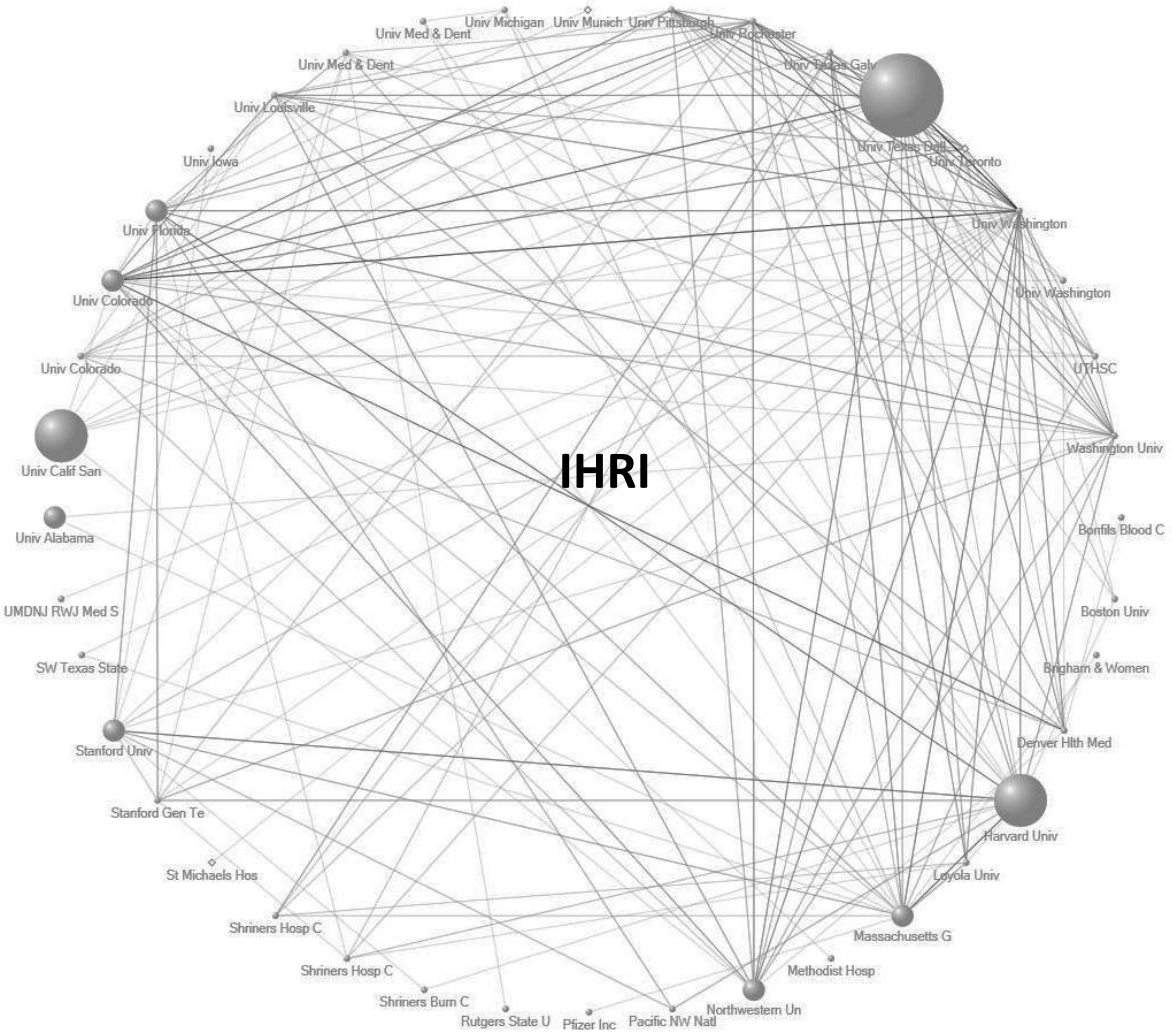
4.3.2 Institutional and Investigator Collaborations

Collaborations among institutions awarded Glue Grant funding were measured by mapping co-author collaborations over the course of the Glue Grant period using a radial network chart **(Figure 4.10)**. Nodes represent author institution, and vertices indicate co-authorship relationship. In addition, we coded the number of Glue Grant principal investigators per institution by the node size and the number of co-authorship relationships by the vertex thickness.

For example, in the CFG consortium, although UC San Diego is the home of the majority of PIs, there are few strong connections with other institutions in the network. Scripps is better integrated into the co-authorship network.




CMC
Univ N Carolina
Univ Nottingham
Univ Oulu
Univ Oxford
Univ Penn
Univ Pittsburgh
Univ Texas
Univ Vermont
Univ Virginia
Univ Wisconsin
UTHSC
Vanderbilt Univ
Weizmann Inst S
Whitehead Inst
Yale Univ
Yeshiva Univ
Ben Gurion Univ
Burnham Inst
CALTECH
Cleveland Clini
Dartmouth Coll
Emory Univ
Florida State U
Harvard Univ
Hebrew Univ Jer
Hubrecht Inst K
Johns Hopkins U
McGill Univ
MIT
Northwestern Un
MRC
NYU
Pacific NW Natl
Scripps Res Ins
Stanford Univ
Stowers Inst Me
Swiss Fed Inst
Technion Israel
Tel Aviv Univ
Thermo Electron
Univ British Co
Univ Calif Berk
Univ Calif Davi
Univ Calif Irvi
Univ Calif Los
Univ Calif San
Univ Calif San
Univ Connecticu
Univ Illinois
Univ Jyvaskyla
Univ Leicester
Univ Leipzig
Univ Liverpool
Univ Med Ctr


IHRI
Univ Michigan
Univ Munich
Univ Pittsburgh
Univ Rochester
Univ Med & Dent
Univ Med & Dent
Univ Texas Galv
Univ Louisville
Univ Texas Bell
Univ Toronto
Univ Iowa
Univ Washington
Univ Florida
Univ Washington
Univ Colorado
UTHSC
Univ Colorado
Washington Univ
Univ Calif San
Univ Alabama
Bonfils Blood C
UMDNJ RWJ Med S
Boston Univ
SW Texas State
Brigham & Women
Stanford Univ
Denver Hlth Med
Stanford Gen Te
Harvard Univ
St Michaels Hos
Loyola Univ
Shriners Hosp C
Massachusetts G
Shriners Hosp C
Methodist Hosp
Shriners Burn C
Northwestern Un
Rutgers State U
Pfizer Inc
Pacific NW Natl

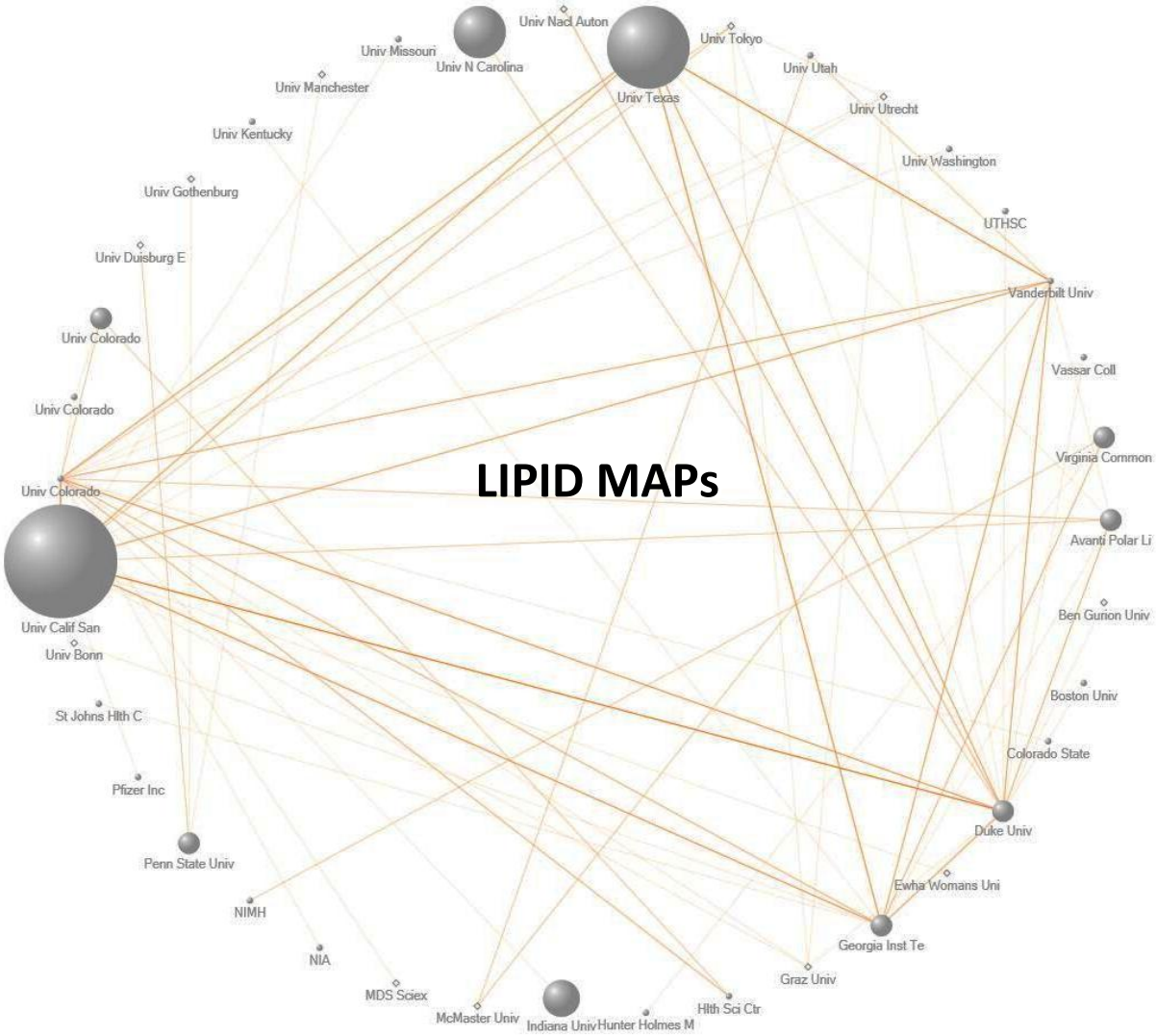


Figure 4.10 Network maps of each consortium, shown for publications during the period 2005-2011. The affiliations around the circle represent all the affiliations appearing in consortium-supported publications, with the size of the node proportional to the number of Glue Grant principal investigators associated with the organization (the smallest size having no PIs). Diamond-shaped nodes represent organizations outside the U.S. The lines represent links by co-authorship in publications, with the thickness of the line proportional to the number of publications (at least two publications must exist between two organizations for a line to exist). SOURCE: Thomson Reuters Web of Science, MEDLINE.

4.3.3 Extent of Interdisciplinary Collaboration

In Section 4.3.1, we discussed the extent to which collaborations, through publication co-authorship, existed within the Glue Grant consortium, among the Glue Grant PIs, and among the institutions on Glue Grant publications. In this section, we examine the nature of the collaborations, specifically the extent of interdisciplinarity.

MeSH terms provide a convenient and accurate way to derive a topical profile of an investigator's publication portfolio. Moreover, quantitative comparisons can be made among different investigators over time, measuring the overlap in MeSH terms used in their respective publications. The goal is to

measure convergence and divergence dynamics of MeSH term overlap among Glue Grant investigators to determine whether any change occurred before and during the Glue Grant period.

4.3.3.1 Scientific Proximity. A calculation known as scientific proximity[16] was used to compare a normalized value (between 0 and 1) for the degree of MeSH descriptor overlap between two investigators. The distance, or proximity, is not symmetric between two researchers. For example, if the scientific proximity between a graduate student and principal investigator were calculated, the overlap of the graduate student to the PI would be perfect or equal to 1.0, assuming all of the student's publications were co-authored by the PI. The PI, on the other hand, will most likely have published a significantly greater number of publications without the student; therefore the PI's overlap would be less than 1.0.

Proximity was measured for each Glue Grant investigator in relation to other investigators within the same consortium for the periods prior to and during Glue Grant periods **(Figure 4.11)**. Each row and column represents one investigator. Investigators are presented in order of increasing number of publications during the period prior to the Glue Grant from left to right and from bottom to top of each matrix. Both the prior and during periods were calculated, and the change is shown in the last panel. The average changes in scientific proximity varied by consortium and were significantly different from zero for all consortia except AFCS.

There are several cases worth closer scrutiny. The PIs in CFG indicated by the arrows 3 and 4 have positive changes in overlap values (+0.37 and +0.47 respectively) with respect to all the other consortium PIs. This suggests that the increase in scientific proximity between these PIs and the other PIs in the consortium is due to a change in their own work between the two periods, rather than drawing other consortium investigators into their own field. Further investigation into these two PIs revealed that PI 3 is a well-established scientist, whose publication output has decreased over time, resulting in fewer, but better overlapped publications with other consortium members. PI 2, however, is a mathematics professor who only recently transitioned to the field of bioinformatics.

The opposite case occurs with PIs illustrated by the arrows 1 and 2 in CMC, whose similarity with the work of others in the consortium has decreased from the prior to during periods. This decrease is largely due to a high similarity at the start, rather than due to a low similarity at the second time point. PI 1 was identified as a crystallographer with a research trajectory that may not correspond with other researchers in the consortium. PI 2 measures the statistics of protein sequences, and is a likely candidate for divergence of research with respect to other consortium members.

[16] http://www.stellman-greene.com/ScientificDistance/Scientific%20Distance%20Report%20Manual.pdf

Investigator Scientific Proximity

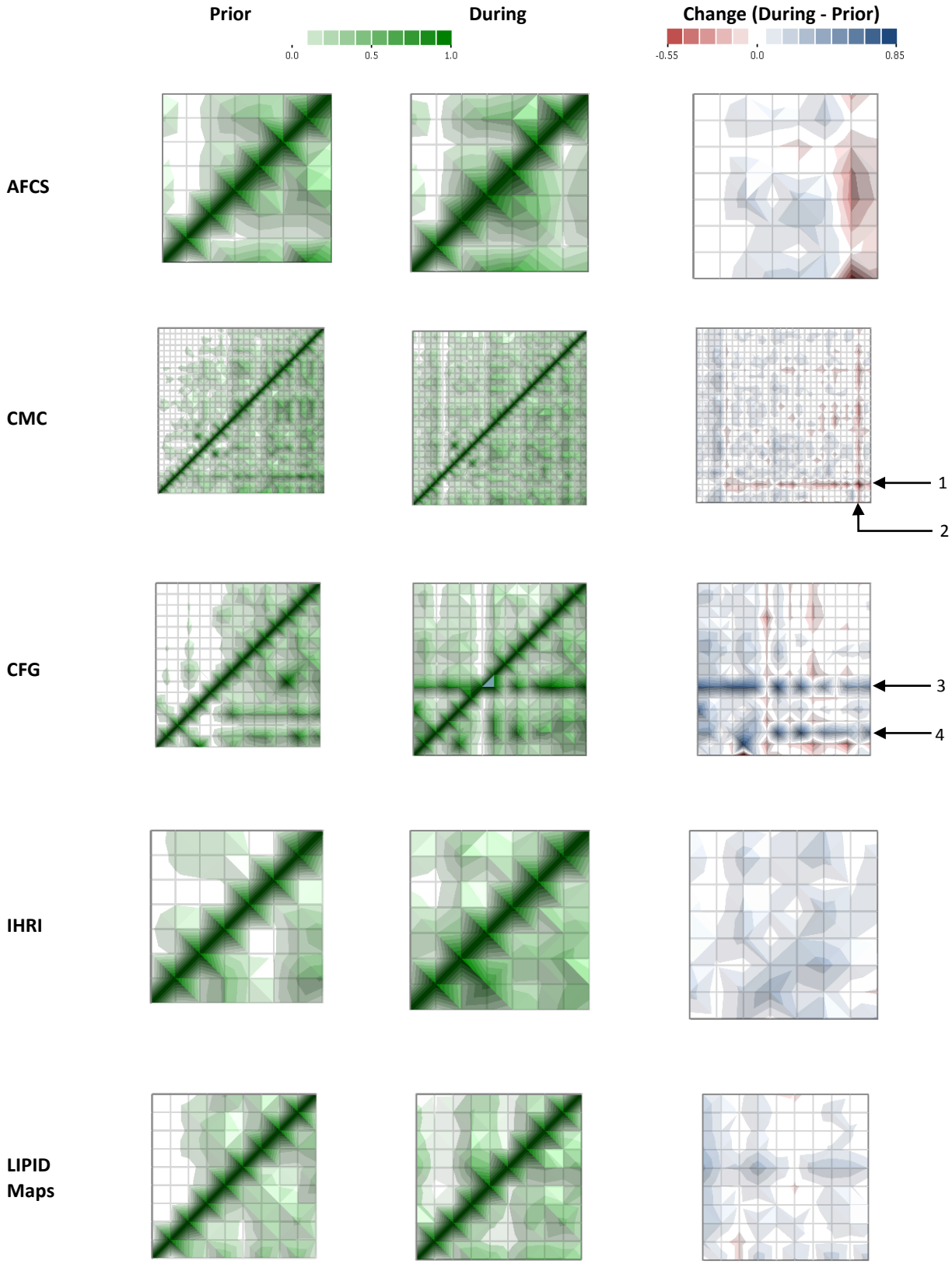


Figure 4.11 PI-by-PI matrix of scientific proximity, defined as the proportion of shared MeSH terms during the period prior to the Glue Grant (1996-2003) and during the Glue Grant (2004-2011). Change in scientific proximity since receiving the Glue Grant compared to prior to the Glue Grant calculated by subtracting the scientific proximity of a given combination of PIs during the period prior to the Glue Grant from that since the Glue Grant was received. Investigators shown are those for whom publications were linked during the prior period and the period during the Glue Grant Investigators are presented in order of increasing number of publications during the period prior to the Glue Grant from left to right and from bottom to top of each matrix. Greater density in scientific proximity is represented as dark green in the first two graphs for each consortium. Greater positive change in scientific proximity is represented as dark blue, while greater negative change in scientific proximity is represented as dark red in the right-most graph for each consortium. SOURCE: MEDLINE

Overall, scientific proximity between investigators significantly increased over time within each consortium (i.e., CMC +0.03, CFG +0.09, IHRI +0.11, LIPID MAPs +0.08, all $p<0.05$) with the exception of investigators in AFCS (+0.01). This suggests that during the Glue Grant period, investigators, despite originating from different fields, collaborated and converged toward one another. Additional details on how the overall change over time was calculated can be found in the Appendix, Section 6.0.

4.3.4 Extent and Impact on Shared Resources

Shared data resources are an increasingly vital feature of large-scale biomedical research. Databases housing gene information, protein structure and function, biomolecular pathways, and other complex interactions are proliferating and becoming crucial to advances in the field. Bioinformatics is a key component of the Glue Grants program, and many in the consortia, particularly the consortia on functional glycomics, cell migration, and LIPID MAPS, were created with the intent that data sharing on IT networks would be implemented to aid both the scientific community surrounding the consortia and the scientific community at large. Although there are many relevant databases covering a variety of molecule classes and research areas, many of the databases did not provide an easy means to link entries to publications. Here we report on articles referencing Glue Grant funding listed in UniProt (a comprehensive protein database) and the Protein Data Bank database housing protein structure information.

4.3.4.1 UniProt Databank Depositions. UniProt collects and houses information on proteins. Entries in the database are centered on unique protein sequences, and cross-references are made to genetic, structural, and functional information contained in other databases. References to publications are entered into the database in one of two ways, (1) manually reviewed information that subject matter experts use to extract useful information from publications, and (2) automatically entered references, typically done by matching protein sequences based in similar genetic information contained in the publication. In **Figure 4.12**, the publication counts per year for manually reviewed references are shown for both Glue Grant and R01 comparison groups. Although the consortia have an overall lower number of references (due in part to fewer publications), the trend suggests that while the number of R01 contributions have decreased, the number of consortium contributions have increased.

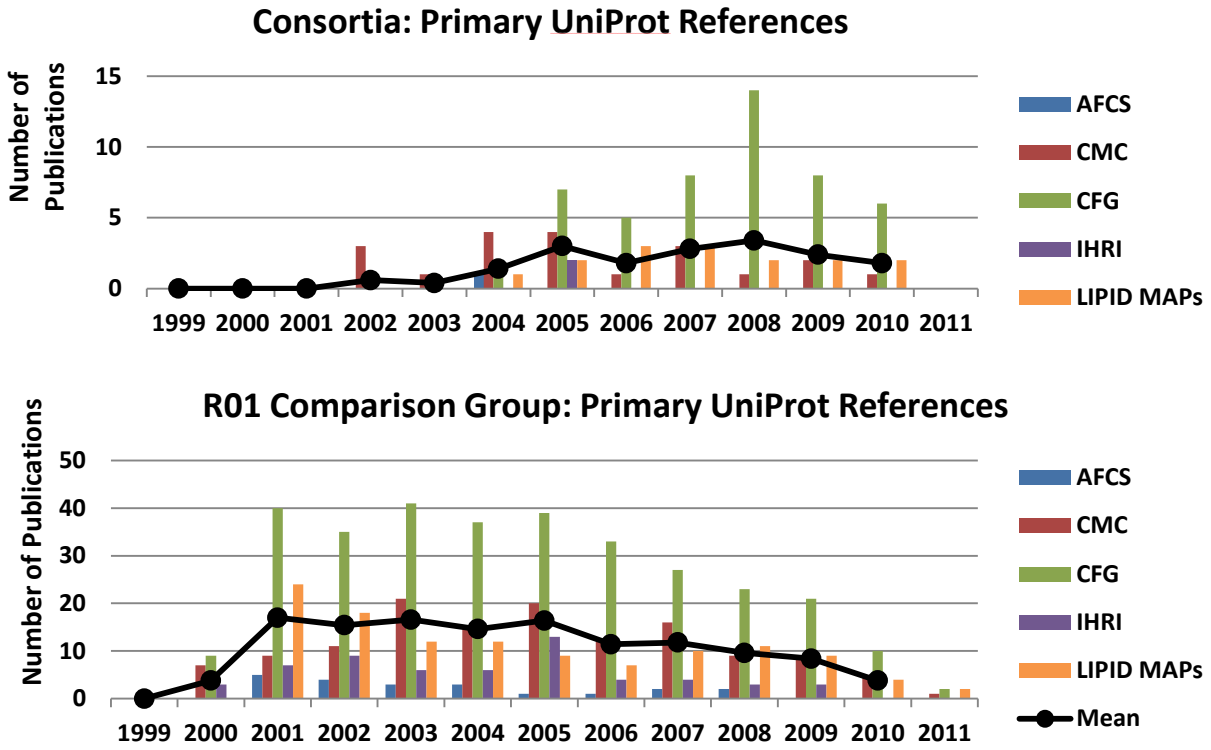


Figure 4.12 Number of publications for the consortium publications (above) as well as their R01 comparison group publications (below) from the years 1999 to 2010 that were referenced as manually curated contributions to the database. The black points and lines represent the mean across the five consortia and their R01 comparisons, respectively. SOURCE: UniProt, MEDLINE.

4.3.3.2 Protein Data Bank Depositions. One of the most widely used repositories is the Protein Data Bank (PDB), which collects and stores structures of proteins. Upon deposition, a publication is typically referenced along with the structure. We accessed the site to identify Glue Grant and R01 publications cited, and found publications from the CMC, CFG, and LIPIS MAPs consortia **(Figure 4.13)**. PDB depositions for CFG and CMC compared to their respective comparison groups are shown in **Figure 4.14**. The data are shown as a percentage of all publications in the group to normalize overall publication counts. The percentage of PDB-linked publications in the CFG and CMC Glue Grant consortia exceeds the respective R01 comparison group.

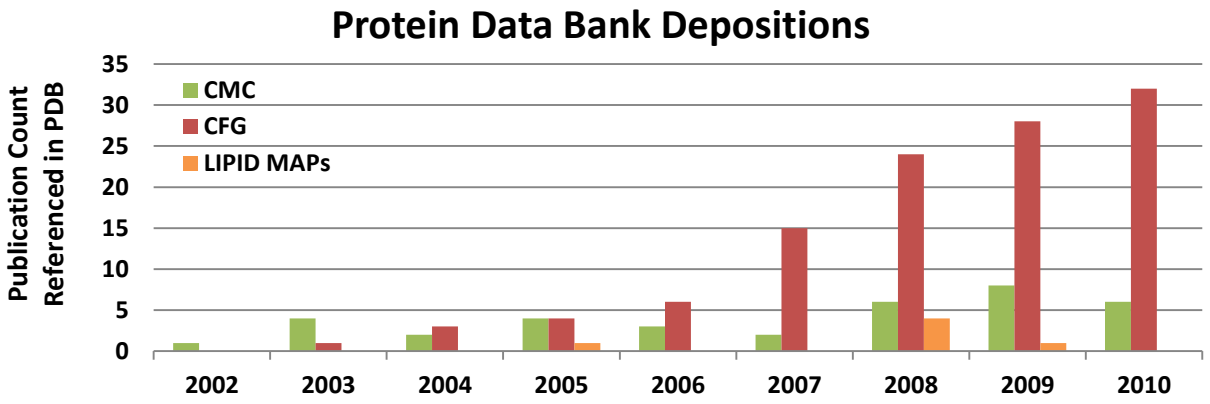


Figure 4.13 Counts of publications referenced in the depositions in the Protein Data Bank (PDB). Medline identification numbers were matched from the PDB data to those identified as acknowledging Glue Grant Support. Three of the five consortia were found to have depositions referenced, all of which are shown above by year. SOURCE: MEDLINE and Protein Data Bank

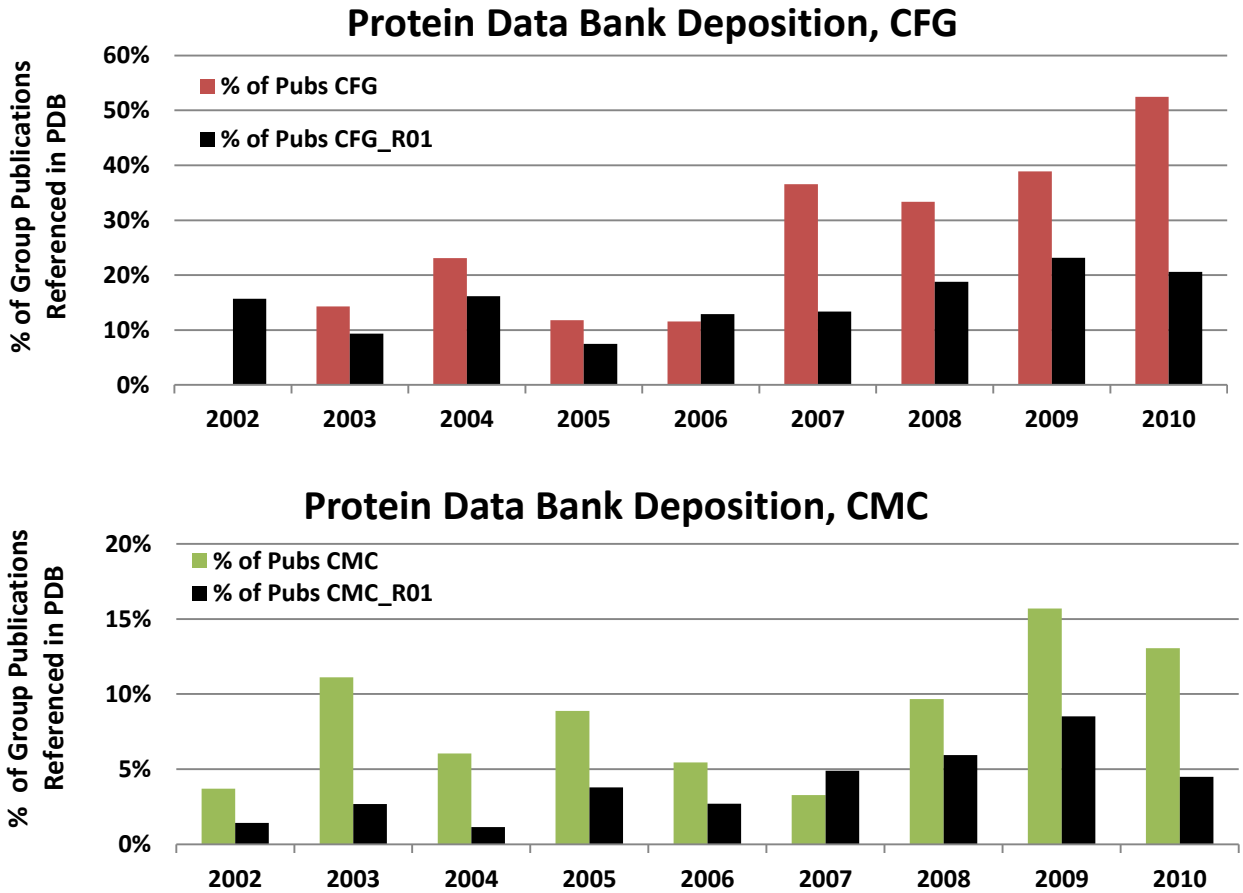


Figure 4.14 Percentage of consortium and comparison group publications referenced in the deposition of protein structures in the Protein Data Bank for the consortia on Functional Glycomics, Cell Migration Group, and their comparison groups. Medline identification numbers were matched from the PDB data to those identified as acknowledging Glue Grant Support. The percentages were calculated for each year shown. SOURCE: MEDLINE and Protein Data Bank

4.3.3.3 Community Use of Consortium Shared Resources. As noted earlier, 24% of publications linked to Glue Grant consortia have no Glue Grant PIs as contributing authors **(Figure 4.15)**. This was most notable for CFG, whose non-PI acknowledgements increased in 2007, corresponding to non-PI papers acknowledged in PDB depositions, indicating that the consortia shared resources were having a broad impact on the research community.

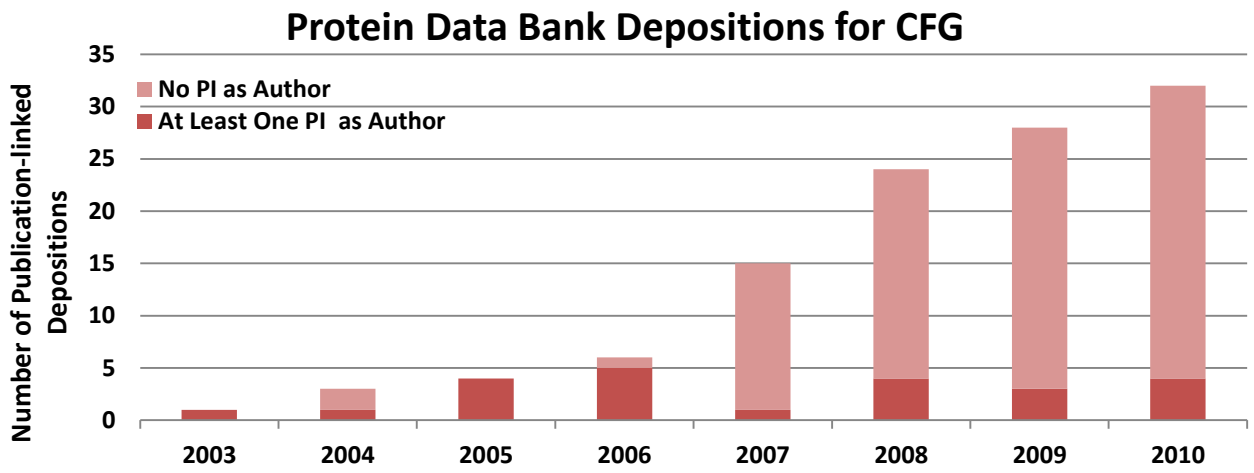


Figure 4.15 Percentage of CFG consortium publications referenced in the deposition of protein structures in the Protein Data Bank. Medline identification numbers were matched from the PDB data to those identified as acknowledging Glue Grant Support. The percentages were calculated for each year shown. SOURCE: MEDLINE and Protein Data Bank

5.0 Conclusions

The goal of this evaluation was to determine the scientific impact of the NIGMS Glue Grants program. We examined a number of outputs, including publications, patents, clinical trials, subsequent funding, databank depositions, PI disciplinarity, and collaboration to assess program impact on investigators and the broader scientific community. We used benchmarks within these measures, such as expected citations and pre/post comparisons, as well as matched comparison groups to benchmark across measures. Overall, we found the Glue Grant program to be meeting its stated goals of stimulating research in topical areas, increasing collaboration, and creating resources and tools for the broader research community. Glue Grant program outputs exceeded that of the R01 benchmarks, indicating that the consortium approach was effective and outperformed topically similar individual grants.

The Glue Grant program has had substantial overall scientific impact when compared to R01 comparison groups, exceeding citation impact per publication over R01 comparison groups, while reaching similar costs per citation within 6 years of the programs. This is the case despite large initial consortium costs as compared with R01 awards.

Publication velocity and productivity among Glue Grant recipients was qualitatively and quantitatively different from R01s funded in the same time period. In almost all cases, productivity reached that of matched R01 awards during the second 5-year award period. Publication citation impact in CMC, AFCS, and CFG was higher than in similar R01 groups.

Additionally, other means of output were observed in particular consortia. Patent awards for new molecules and scientific uses were identified for two consortia (e.g., CMC, AFCS), although the other consortia have patent applications in process. IHRI met one of their stated goals for translational research by concluding two clinical trials during the Glue Grant period.

The Glue Grants fostered increased scientific interaction among consortium investigators and increased interactions with the outside scientific community.

All Glue Grant investigators had more high-author count publications during the Glue Grant period than before, as well as greater median values than their R01 comparison groups. Glue Grant publications in the LIPID MAPs, AFCS, and IHRI groups featured more authors per publication than for publications by the same investigators outside the Glue Grant. Surprisingly, 24% of publications acknowledging Glue Grant support featured no PIs as authors, the majority of which were identified as core facility users. Additionally, 23% of all Glue Grant publications featured more than one PI.

Shared data repository activity was a primary output for CMC and CFG groups, as observed by references to UniProt and the Protein Data Bank. These depositions were a greater part of the consortia activity than for the R01, as determined by percentage of publications referenced.

Glue Grant investigators from fields outside the pure biomedical sciences were observed to publish in more interdisciplinary publications during the Glue Grant period, as observed by overlap of MeSH terms with other Glue Grant investigators. In addition, the MeSH term overlap used to calculate scientific proximities of individual Glue Grant investigators enabled us to identify investigators whose scientific background differed from most of the other investigators in the consortia, as well as determine the extent to which their scientific paths converged. Overall, scientific proximity between investigators significantly increased over time within each consortium with the exception of investigators in AFCS (e.g., CMC, CFG, IHRI, LIPID MAPs).

The Glue Grant program was designed to encourage investigators to work together to solve known biomedical research problems in ways that investigator-focused R01 awards could not facilitate. Publication and shared repository data indicate that consortia activities did indeed facilitate multidisciplinary research, engagement with the outside scientific community, and highly collaborative scientific work. The structure of the Glue Grants produced clear collaborative networks not seen in similar R01s. However, the high productivity of these researchers, even prior to the Glue Grant period, limited our ability to assess to what extent the Glue Grants accelerated investigator productivity and impact. .

Investigators participating in the Glue Grant program were able, to some degree, to leverage their participation to win new NIH awards.

Glue Grant investigators were observed to be well funded prior to the Glue Grant when compared with similar investigators in the comparison groups. Although there is no clear evidence that the Glue Grants directly resulted in more and higher valued research awards by investigators, the data suggest that investigators in all consortia except AFCS received increased funding during the Glue Grant period; particularly CFG. Additionally, co-funding outside the Glue Grants occurred between two investigators in CMC beginning in 2009.

The data repositories created on the individual consortium websites were difficult to access and difficult to link to publications and individual contributors.

A primary programmatic goal of the consortia was to standardize and establish consortia-led data repositories for use by the greater scientific community. Although such repositories were created, their impact could not be evaluated because they had insufficient accessibility. Nor were clear references to publications or investigators available in an easily accessible data format that would allow evaluation. It is recommended that Medline identification numbers, investigator names, and organization names be included in deposition events for evaluation and tracking of use.

Despite the focus of the consortia to establish collaborative research, it was unclear to what extent collaborations took place, other than co-authorship of publications acknowledging Glue Grant support.

Collaborative research in this study was measured by publications co-authored by cross-institutional and multi-PI co-authors on publications acknowledging Glue Grant support. There was evidence that both a greater number of authors on average contributed to Glue Grant publications than R01 comparison groups and that greater collaboration across institutions took place. However, information collected from grantees on collaborations with other principal investigators not directly receiving Glue Grant support would greatly facilitate studies of this nature. This would provide a means by which to evaluate an additional group of PIs whose research was supported by the Glue Grants.

6.0 Appendix: Supporting Information

6.1 Glue Grant Investigators

Glue grant investigators are defined as individuals whose names are listed as the principal investigator on any funded grant under one of five Glue Grant project numbers, represented by each of the five consortia. The number of PIs is shown in **Figure 6.1** and the list of individuals is shown in **Table 6.1**.

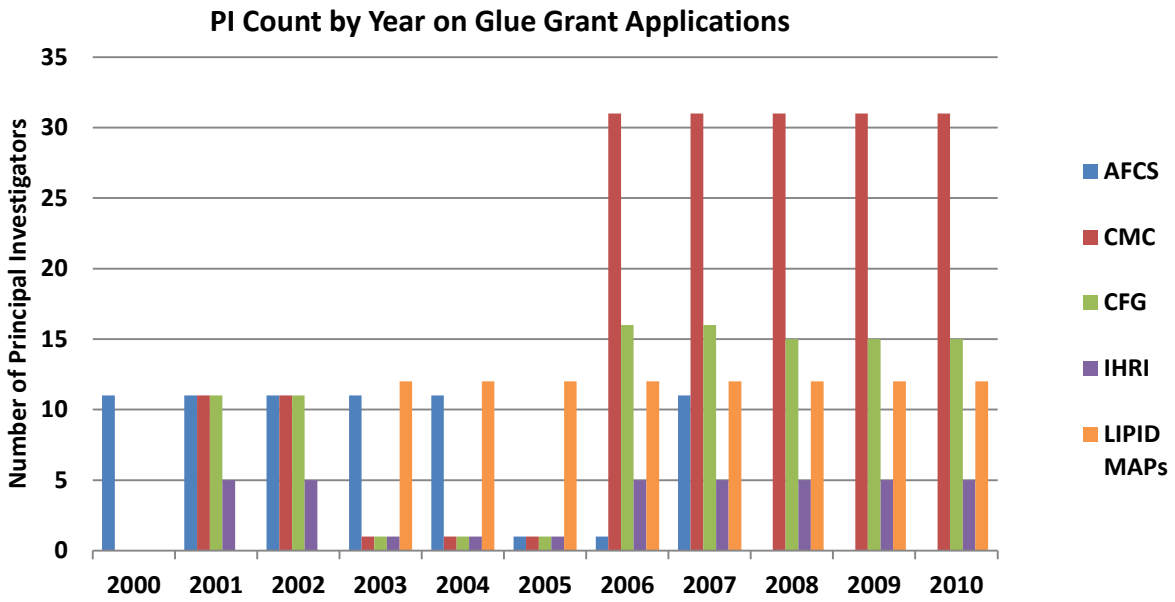


Figure 6.1 Number of Glue Grant principal investigators in the present study, by year and consortium. Although the number of investigators varied by year, only the unique total for each consortium was used to normalize the data, unless otherwise noted in the text or figure. SOURCE: NIH RePORTer and Glue Grant Progress Reports

Table 6.1 Glue Grant Investigators

Last Name	First Name	Middle Name	Consortium	# of Funded Applications[17]
HORWITZ	ALAN	F	CMC	29
GINSBERG	MARK	HOWARD	CMC	8
HUNT	DONALD	F	CMC	8
JACOBSON	KENNETH	A	CMC	8
LAUFFENBURGER	DOUGLAS	A	CMC	8
LIDDINGTON	ROBERT		CMC	8
PEARSON	WILLIAM	R	CMC	8
SCHWARTZ	MARTIN	A.	CMC	8
TAYLOR	KENNETH	ALLEN	CMC	8

[17] The "# of funded applications" column indicates the count of individually funded Glue Grant applications associated with a principal investigator. The count includes all sub-projects that are listed under the primary Glue Grant project number.

BRUGGE	JOAN	SIEFERT	CMC	5
BURRIDGE	KEITH		CMC	5
CAMPBELL	IAIN	D	CMC	5
CONDEELIS	JOHN	S	CMC	5
DANUSER	GAUDENZ		CMC	5
FOX	JAY	WILLIAM	CMC	5
GEIGER	BENNY		CMC	5
GRATTON	ENRICO		CMC	5
HAHN	KLAUS	MICHAEL	CMC	5
HANEIN	DORIT		CMC	5
HANNAY	TIMO		CMC	5
HAUGH	JASON	M.	CMC	5
IMPERIALI	BARBARA		CMC	5
KLEMKE	RICHARD	L.	CMC	5
LOEW	LESLIE	M	CMC	5
MOGILNER	ALEXANDER		CMC	5
MONTELL	DENISE	J.	CMC	5
PARSONS	JOHN	T	CMC	5
SCHWARZBAUER	JEAN	E	CMC	5
SLIGAR	STEPHEN	G.	CMC	5
VALE	RONALD	D	CMC	5
VOLKMANN	NIELS		CMC	5
GRIFFITH	LINDA	G.	CMC	3
HYNES	RICHARD	O	CMC	3
MAYER	TOBIAS		AFCS	22
GILMAN	ALFRED	G	AFCS	16
GRAY	MARTHA	L	AFCS	11
MICHNICK	STEPHEN		AFCS	11
MUMBY	SUSANNE	M	AFCS	11
OLSON	ERIC	N	AFCS	11
SEAMAN	WILLIAM	E	AFCS	11
SIMON	MELVIN		AFCS	11
STERNWEIS	PAUL	C	AFCS	11
SUBRAMANIAM	SHANKER		AFCS	11
TSIEN	ROGER	Y	AFCS	11
PAULSON	JAMES	C.	CFG	18
HEAD	STEVEN	R	CFG	8
WILSON	IAN	A	CFG	8
ALVAREZ	RICHARD		CFG	5
BLIXT	OLA	KLAS	CFG	5
CUMMINGS	RICHARD	D	CFG	5
DELL	ANNE		CFG	5

FINN	M.G.		CFG	5
GOLDBERG	DAVID		CFG	5
KANEHISA	MINORU		CFG	5
KNIREL	YURLY		CFG	5
NOVOTNY	MILOS		CFG	5
ORR	SALLY	L	CFG	5
SASISEKHARAN	RAM		CFG	5
SOBIESZCZUK	PETER		CFG	5
VAN KOOYK	YVETTE		CFG	5
CROCKER	PAUL	R	CFG	3
HEDRICK	STEPHEN	M	CFG	3
KRONENBERG	MITCHELL		CFG	3
LE	DZUNG	T	CFG	3
LOWE	JOHN	B.	CFG	3
VAN HALBEEK	HERMAN		CFG	3
VENKATARAMAN	GANESH		CFG	3
TOMPKINS	RONALD	GARY	IHRI	40
DAVIS	RONALD	WAYNE	IHRI	8
MOLDAWER	LYLE	L	IHRI	5
SCHOENFELD	DAVID	ALAN	IHRI	5
SMITH	RICHARD	D	IHRI	5
CHAUDRY	IRSHAD	H	IHRI	3
MAIER	RONALD	VITT	IHRI	3
WEST	MICHAEL	ALLAN	IHRI	3
DENNIS	EDWARD	A	LIPID MAPS	30
GLASS	CHRISTOPHER	K	LIPID MAPS	16
MURPHY	ROBERT	CARL	LIPID MAPS	16
SUBRAMANIAM	SHANKAR		LIPID MAPS	16
BROWN	ALEX		LIPID MAPS	8
MERRILL	ALFRED	HARRISON	LIPID MAPS	8
RAETZ	CHRISTIAN	R	LIPID MAPS	8
RUSSELL	DAVID	W.	LIPID MAPS	8
SHAW	WALTER	A	LIPID MAPS	8
VANNIEUWENHZE	MICHAEL	S	LIPID MAPS	8
WITZTUM	JOSEPH	L.	LIPID MAPS	8
WHITE	STEPHEN	H.	LIPID MAPS	5
WINOGRAD	NICHOLAS		LIPID MAPS	3

6.2 Collaboration Statistics for Glue Grant Investigators

The following tables present collaboration statistics for Glue Grant investigators. For Tables 6.2 through 6.6, "Total" is calculated for all publications during the period from 1996-2011, "Prior" indicates the period before the Glue Grants (defined as 1996-2003), "Grant" represents all Glue Grant publications, and "During" refers to non-Glue Grant publications during the period "2004-2011."

Table 6.2 Collaboration Metrics for AFCS Glue Grant Investigators[18]

	Number of Publications				Median Author Count				Unique Coauthors				New Coauthors	
GGPI	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Grant**	**During**
1861865	55	51	0	4	4	4	--	2	31	30	0	1	0	1
1887103	98	49	4	45	5	5	4	5	39	21	4	20	2	17
1888540	25	12	5	8	7	5	9	7.5	14	6	2	8	2	6
1898416	34	19	7	8	5.5	5	11	5.5	20	15	5	4	5	0
1898994	161	46	0	115	6	6	--	6	44	22	0	28	0	22
1919810	8	5	1	2	6.5	4	11	8	8	5	2	3	2	1
1969499	14	6	1	7	4.5	3.5	4	6	4	2	1	3	0	2
6060863	1	1	0	0	2	2	--	--	0	0	0	0	0	0
6230993	52	32	11	9	5	4	9	7	45	23	5	22	5	17
6727026	2	0	2	0	4	--	4	--	1	0	1	0	1	0
6727037	1	0	1	0	29	--	29	--	2	0	2	0	2	0
Average	**41.0**	**20.1**	**2.9**	**18.0**	**7.1**	**4.3**	**10.1**	**5.9**	**18.9**	**11.3**	**2.0**	**8.1**	**1.7**	**6.0**
p-value			***0.02***	0.19			***0.04***	0.23			***0.02***	0.07		0.12

[18] 'Total' publications refer to all publications from 1994-2011 co-authored by a Glue Grant investigator, linked by name; 'Prior' is defined to the period 1994-2003, while 'During' is defined as the period following that, 2004-2011. 'Grant' publications are the publications acknowledging support from Glue Grants co-authored by that particular investigator. Unique co-authors are the counts of unique co-authors appearing in publications within each publication subset. New coauthors are the coauthors in 'Grant' or 'During' publications that do not appear in the 'Prior' group. These definitions apply for Tables 6.2 to 6.6.

Table 6.3 Collaboration Metrics for CMC Glue Grant Investigators

	Number of Publications				Median Author Count				Unique Coauthors				New Coauthors	
GGPI	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Grant**	**During**
1840203	106	37	5	64	5	4	5	5	34	13	5	24	4	18
1860230	236	79	19	138	6	6	4	6	61	29	11	32	10	23
1860713	97	32	7	58	6	5.5	10	6	49	31	4	22	2	16
1863307	61	29	4	28	4	4	4	5	38	24	6	18	3	12
1863319	120	53	2	65	4	4	9	4	43	27	0	22	0	16
1863912	145	78	15	52	5	5	5	5	102	73	6	51	5	26
1864076	123	59	9	55	6	5	7	5	67	32	8	40	6	34
1865701	51	18	21	12	5	5	5	5	12	9	1	2	1	2
1869149	37	13	5	19	3	2.5	3.5	3	6	1	1	5	1	4
1874802	28	12	4	12	4.5	3.5	6.5	5.5	10	4	1	7	1	6
1882706	18	11	0	7	3	3	--	2	6	5	0	1	0	1
1883049	78	20	21	37	3	2	3	3.5	18	3	6	11	6	10
1887665	83	44	6	33	5	4	5.5	5	43	23	3	21	3	17
1893068	28	11	5	12	5	2	5	6	11	2	3	8	3	7
1893792	38	16	9	13	5	4	4	8	16	3	5	12	4	11
1894213	110	72	1	37	6	6	6	6	59	42	0	19	0	17
1895083	82	32	2	48	4	4	2	4	34	23	0	12	0	11
1895514	40	15	3	22	5.5	5	8	5	7	1	4	3	4	3
1901683	168	66	7	95	8	9	6	8	116	49	8	76	8	64
1965211	47	21	13	13	4	4	5	5	23	6	12	8	11	6
2165137	75	35	1	39	4	3	5	5	25	14	2	13	2	9
2211633	66	19	4	43	6	5	6.5	7	46	15	4	32	4	27
3117594	10	2	5	3	6	6	5	5	8	3	5	4	5	2
3131262	43	7	7	29	6	2.5	8	6	24	2	4	21	4	20
6271603	23	2	5	16	2	1.5	2	3	1	1	0	0	0	0
6407777	28	11	4	13	6	3	7	5.5	17	8	7	11	3	6
6620738	57	9	11	37	5	4	5	5	20	3	5	17	4	14
6768673	42	4	19	19	4	2.5	4	4	14	2	11	3	10	2
6799074	28	8	4	16	6	5	6	7	22	8	4	19	0	14
8171084	52	22	2	28	3	3	2.5	3	24	9	0	16	0	15
8815636	1	1	0	0	5	5	--	--	1	1	0	0	0	0
8815796	20	0	2	18	6	--	7.5	2.5	3	0	1	2	1	2
Average	**66.9**	**26.2**	**6.9**	**33.8**	**4.9**	**4.1**	**5.4**	**5.0**	**30.0**	**14.6**	**4.0**	**16.6**	**3.3**	**13.0**
p-value			***<0.0001***	***<0.0001***			***0.004***	0.42			***0.00***	***0.0001***		***0.0001***

Table 6.4 Collaboration Metrics for CFG Glue Grant Investigators

	Number of Publications				Median Author Count				Unique Coauthors				New Coauthors	
GGPI	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Grant***	**During***
1863380	100	39	0	61	4	3	--	5	20	3	0	19	0	17
1864103	144	66	3	75	7	7	9	7	89	55	7	48	5	32
1867470	104	50	7	47	6	5	8	6	29	14	4	18	3	15
1878922	323	100	23	200	8	7	7.5	10	135	54	33	84	27	61
1882172	39	17	0	22	5	5		5	29	15	0	15	0	14
1884718	120	53	28	39	5	4	7	5	43	20	15	17	13	11
1919410	1	0	1	0	20	--	20	--	2	0	2	0	2	0
1922297	50	5	3	42	5	5	6	4.5	21	6	1	17	1	14
1965189	52	11	18	23	5	5	6.5	5	29	5	17	14	16	13
2089285	25	3	17	5	8	6	8	8.5	27	3	21	7	20	7
6348123	10	3	5	2	7.5	14	6	2	13	7	5	2	5	2
6510825	9	8	1	0	5	5.5	5	--	4	3	2	0	1	0
6616849	53	21	14	18	9	8	9	9	18	7	7	8	7	4
7007911	13	3	6	4	5	5	6	5	6	1	6	1	5	0
7007999	23	1	17	5	8	4	8	8.5	25	2	21	7	19	6
8319700	57	34	10	13	7	7	5	8	30	21	7	11	3	6
8728766	22	0	20	2	7	--	7	4	16	0	16	3	16	3
8817799	1	0	1	0	20	--	20	--	2	0	2	0	2	0
8817812	14	0	14	0	8.5	--	8.5	--	15	0	15	0	15	0
8817826	25	0	3	22	7	7	8	6	6	1	6	0	5	0
10297254	1	1	0	0	5	5	--	--	1	1	0	0	0	0
Average	**56.5**	**19.8**	**9.1**	**27.6**	**7.7**	**6.0**	**8.6**	**6.2**	**26.7**	**10.4**	**8.9**	**12.9**	**7.9**	**9.8**
p-value			0.08	0.07			0.40	0.12			0.66	0.30		0.51

Table 6.5 Collaboration Metrics for IHRI Glue Grant Investigators

	Number of Publications				Median Author Count				Unique Coauthors				New Coauthors	
GGPI	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Grant**	**During**
1878566	155	85	17	53	8	7.0	12.0	8.0	76	27	32	31	29	24
1881995	121	46	17	58	7	7.0	12.0	7.0	56	18	16	29	14	26
1883223	37	20	10	7	5	4.5	13.0	3.0	11	0	11	0	11	0
1898321	97	41	28	28	6	5.0	13.0	5.0	48	18	25	16	22	9
2090608	93	40	26	27	7	6.0	10.5	6.0	57	27	26	15	25	10
2419504	291	147	3	141	6	5.0	14.0	7.0	22	6	8	11	8	9
7508231	29	8	9	12	8	6.0	11.0	8.5	50	13	15	24	15	23
8080536	306	81	14	211	7	6.0	9.0	7.0	77	12	9	63	8	58
Average	**15.5**	**58.5**	**15.5**	**67.1**	**6.8**	**5.8**	**11.8**	**6.4**	**49.6**	**15.1**	**17.8**	**23.6**	**16.5**	**19.9**
p-value			***0.04***	0.09			***0.0001***	***0.001***			0.17	0.46		0.67

Table 6.6 Collaboration Metrics for LIPID Glue Grant Investigators

	Number of Publications				Median Author Count				Unique Coauthors				New Coauthors	
GGPI	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Total**	**Prior**	**Grant**	**During**	**Grant**	**During**
1858908	135	64	39	32	4.0	4.0	6.0	4.0	49	24	20	13	18	7
1863389	106	54	23	29	5.0	4.0	9.0	6.0	38	18	17	9	13	7
1871016	88	39	31	18	7.0	5.0	9.0	7.0	60	27	29	18	23	11
1878773	94	40	15	39	7.0	8.0	11.0	6.0	46	23	14	22	12	12
1881159	123	56	31	36	4.0	4.0	5.0	4.0	59	23	15	32	14	26
1897859	51	5	4	42	5.0	5.0	4.5	4.5	7	2	1	7	0	5
1898342	63	44	11	8	5.0	5.0	10.0	4.0	31	20	9	2	9	2
1910867	19	0	13	6	6.0	--	6.0	4.0	20	0	15	5	15	5
1921538	8	0	2	6	3.0	--	10.5	2.5	1	0	0	1	0	1
1936703	46	1	22	23	5.0	5.0	7.0	4.0	23	0	15	11	15	11
1962547	134	67	6	61	7.0	6.0	11.0	8.0	94	58	6	59	0	36
1964913	63	29	4	30	3.0	3.0	3.0	3.5	14	8	0	7	0	6
6597841	4	0	4	0	18.5	--	18.8	--	5	0	5	0	5	0
Average	**71.8**	**30.7**	**15.8**	**25.4**	**6.1**	**4.9**	**8.5**	**4.8**	**34.4**	**15.6**	**11.2**	**14.3**	**9.5**	**9.9**
p-value			***0.03***	0.11			***0.003***	***0.002***			0.36	0.53		0.92

6.3 Scientific Proximity Measurement

Scientific proximity is a measure of the similarity between the work of two PIs and is calculated as:

$$Proximity_{i,j} = \frac{Number\ of\ MeSH\ Terms\ Shared_{i,j}}{Total\ number\ of\ MeSH\ Terms_i} \qquad \textbf{Equation 1}$$

This number ranges from 0 to 1, where 1 is complete overlap between MeSH terms associated with the two PIs, and 0 is complete discordance between the two PIs. Proximity is asymmetric, so that for a comparison between the overlap in MeSH terms for the same two PIs, an alternative calculation is:

$$Proximity_{j,i} = \frac{Number\ of\ MeSH\ Terms\ Shared_{i,j}}{Total\ number\ of\ MeSH\ Terms_j} \qquad \textbf{Equation 2}$$

The key difference is that the denominator for **Equation 1** is that of the total number of MeSH terms associated with PI_i while that for **Equation 2** is the total number of MeSH terms associated with PI_j.

Change in proximity between two PIs can be calculated by partitioning publications across time and subtracting the calculations made within the given time periods. This change, when averaged together for a given PI, may indicate this PI's ability to draw other into his/her field (Θ_1) or a change in the PI's own work to reflect greater similarity with work of other investigators (Θ_2). A measure of the former for PI_x in a field with *n* investigators comparing times 1 (t_1) and 2 (t_2), can be represented by:

$$\theta_{1,x} = \frac{1}{n-1}\sum_{i=1}^{n}\left[(Proximity_{x,i})_{t_2} - (Proximity_{x,i})_{t_1}\right] \qquad \textbf{Equation 3}$$

while the latter may be represented by:

$$\theta_{2,x} = \frac{1}{n-1}\sum_{i=1}^{n}\left[(Proximity_{i,x})_{t_2} - (Proximity_{i,x})_{t_1}\right] \qquad \textbf{Equation 4}$$

The key feature that is different for Θ_1 is that the denominator for a given PI_x does not change within a given time period and is the total number of MeSH terms associated with PI_x during that time period. In contrast, the denominator in each term within the summation of Θ_2 for $PI_x$is different for each term within the summation and is the total number of MeSH terms associated with PI_i within the time period.

For the data presented in **Figure 4.11**, Θ_1 is shown horizontally, while Θ_2 is shown vertically. For example, indicated by the arrows 1 and 2, both PIs have high values for Θ_1 (0.37 and 0.47 respectively). This suggests that the increase in scientific proximity between these PIs and the other PIs in the consortia is due to a change in their own work, rather than the ability of these individuals to draw others in the consortia into their own field. This is consistent with the fact that they may be less well established than other PIs in the consortia, with fewer publications in the prior period than other (26 and 70 during the period prior, 51 and 7 during the grant period respectively, versus median overall 124 during the prior period and 195 during the grant period). As an opposing example, arrow 3 indicates a PI for whom similarity between his/her own work and the work of others in the consortia has decreased from time 1 to time 2. Although this decrease is startling, it is largely due to a high similarity at the start, rather than due to a low similarity at time 2.

The average change in scientific proximity varied by consortium, and was statistically significantly different from zero for all consortia except AFCS. Averaged across PIs, mean $\Theta_1$and Θ_2 will be

mathematically equivalent. However, their standard deviation will be different, with corresponding differences in statistical significance **(Table 6.7)**.

Table 6.7 Mean Θ and statistical significance by type of change in proximity

Consortium	Mean Θ		SD	p
AFCS	+0.01	Θ_1	0.04	0.64
		Θ_2	0.12	0.86
CMC	+0.03	Θ_1	0.05	***0.005***
		Θ_2	0.07	***0.04***
CFG	+0.09	Θ_1	0.14	***0.02***
		Θ_2	0.11	***0.007***
IHRI	+0.11	Θ_1	0.04	***<0.0001***
		Θ_2	0.06	***0.002***
LIPID MAPs	+0.08	Θ_1	0.06	***0.002***
		Θ_2	0.07	***0.008***

By both measures, PIs within each consortia became more similar to one another after the Glue Grant compared to before the Glue Grant than would be expected among a random group of PIs (i.e., mean Θ = 0), with the exception of the PIs in the AFCS consortium.