

出國報告（出國類別：其它）

出席 2012 全球生物多樣性資訊學研  
討會  
(Global Biodiversity Informatics  
Conference, GBIC)

會議報告

服務機關：農委會林業試驗所

姓名職稱：王豫煌特聘研究員

派赴國家：丹麥

出國期間：101 年 7 月 1 日至 7 月 6 日

報告日期：101 年 9 月 18 日

# 目 錄

1. 會議期間 .....	1
2. 會議地點 .....	1
3. 會議網址 .....	1
4. 會議目的 .....	1
5. 研討會議程 .....	1
6 會議重要內容記述 .....	5
6.1 主題演講.....	5
6.2 分組討論與結論.....	6
7 會議最後的重點結論： .....	6
8 附錄一主題演講簡報內容 .....	8
9 會議活動照片 .....	45

# 2012 全球生物多樣性資訊學研討會 (Global Biodiversity Informatics Conference, GBIC) 會議報告

## 1. 會議期間

2012 年 7 月 1-4 日

## 2. 會議地點

丹麥哥本哈根

## 3. 會議網址

<http://www.gbic2012.org/index.html>

## 4. 會議目的

此會議由全球生物多樣性資訊機構 (Global Biodiversity Informatics Facility, GBIF)、生物多樣性研究資訊基礎設施與活動國際協調組織 (Coordination of Research e-Infrastructures Activities Toward an International Environment for Biodiversity, CReATIVE-B)、倫敦自然史博物館 (Natural History Museum, London)、生命大百科 (Encyclopedia of Life, EOL)、生命條碼聯盟 (Consortium for the Barcode of Life, CBOL) 等組織共同舉辦。目的在於探討如何以生物學、電腦資訊科學及社群網路的最新發展趨勢和技術來解決全球生物多樣性不斷快速流失和生物資源管理的危機。會議主辦單位邀集 100 多位來自各國生物、生態、電腦資訊科學、官方政策等各領域的科學家，分組進行三天密集的腦力激盪，討論愛知目標 (Aichi Targets) 重大的生物多樣性議題及其應用整合科學方法的可能解決之道。分組討論過程中所有科學家提出的想法、建議，再經過綜合的整理和論述，將集結成“全球生物多樣性資訊學展望 (Global Biodiversity Informatics Outlook, GBIO)”做為未來十年達成阻止全球生物多樣性持續損失的跨領域整合科學方法參考指引，以支援保育優先策略擬定和政策執行。

## 5. 研討會議程

2012 年 7 月 1 日

19:00-21:00

晚間聚會 哥本哈根大學動物博物館(Zoological Museum)

第一日： 2012 年 7 月 2 日

8:45-9:00	開幕式及會議介紹		
9:00-10:20	主題演講： <ul style="list-style-type: none"> <li>• Lucas Joppa: "Science-driven data"</li> <li>• Daan du Toit: "The policy - research infrastructure interface: implications for biodiversity informatics"</li> </ul>		
10:20-10:50	Coffee break		
10:50-11:30	主題演講： <ul style="list-style-type: none"> <li>• Robert Robbins: "Half of Our Biosphere is (No Longer) Missing: Implications for Understanding Biodiversity, Bioinformatics, and Life Itself"</li> </ul>		
11:30-12:30	GBIC 團隊歡迎與介紹		
	紅隊主席群 <ul style="list-style-type: none"> <li>• Cyndy Parr</li> <li>• Alberto Apostolico</li> <li>• Jeff Price</li> <li>• Jerry Harrison</li> </ul>	綠隊主席群 <ul style="list-style-type: none"> <li>• Elizabeth Arnaud</li> <li>• Rod Page</li> <li>• Selwyn Willoughby</li> <li>• Gregoire Dubois</li> </ul>	藍隊主席群 <ul style="list-style-type: none"> <li>• Dawn Field</li> <li>• Bryan Heidorn</li> <li>• Dora Canhos</li> <li>• Juan Bello</li> </ul>
12:30-1:30	午餐		
1:30-3:00	第一日 平行分組討論： 生物多樣性科學－資訊學和資料管理如何增進我們對生物多樣性的瞭解、如何支援生物多樣性的基礎與應用研究？		



	Biodiversity Science - How can informatics and data management support better understanding of biodiversity and support basic and applied biodiversity research?
3:00-3:30	休息
3:30-5:30	第一日 平行分組討論 (繼續)
7:00-10:00	晚宴
<b>第二日： 2012 年 7 月 3 日</b>	
8:30-9:00	第一日平行分組討論成果之整合摘述
9:00-10:00	<p>第二日 平行分組討論：</p> <p>愛知目標 (全球政策優先目標之摘要)</p> <p>資訊學和資料管理如何支援達成這些目標所需的資訊需求與全球行動？</p> <p>Aichi Targets (as summary of global policy priorities)</p> <p>- How could informatics and data management support the information needs and global response to these targets?</p>
10:00-10:30	Coffee break
10:30-12:30	第二日 平行分組討論 (繼續)
12:30-1:30	Lunch
1:30-3:00	第二日 平行分組討論 (繼續)
3:00-3:30	Coffee break

3:30-5:30	第二日 平行分組討論 (繼續)
第三日： 2012 年 7 月 4 日	
8:30-9:00	第二日平行分組討論成果之整合摘述
9:00-10:00	<p>第三日 平行分組討論：</p> <p>建立視野－兩日的平行分組討論關於資訊學的需求和機會，各團隊對於以下各項生物多樣性議題的建議：</p> <ul style="list-style-type: none"> <li>• 內容需求</li> <li>• 資訊學需求</li> <li>• 文化變革需求</li> </ul> <p>Establishing the Vision - From the discussions of informatics needs and opportunities in Day 1 and Day 2, what are the recommendations from each team for:</p> <ul style="list-style-type: none"> <li>• <b>Content needs</b></li> <li>• <b>Informatics needs</b></li> <li>• <b>Cultural change needs</b></li> </ul>
10:00-10:30	休息
10:30-12:30	第三日 平行分組討論 (繼續)
12:30-1:30	午餐
1:30-3:00	第三日平行分組討論成果之整合摘述
3:00-3:30	休息
3:30-5:30	<p>GBIC 成果摘述與下一步發展之討論</p> <p>Summary of GBIC outcomes and discussion of next steps</p>

## 6 會議重要內容記述

### 6.1 主題演講

#### 6.1.1 科學研究導向的資料收集 (Science-driven data)

過去累積了很多資料，形成資料洪流 (data deluge) 而塑造資料導向的科學研究 (data-driven science)；然而，此現象是否真能解決全球所面臨的生物多樣性不斷快速損失的難題？儘管我們現在看起來像是被淹沒在資料洪流中，但實際狀況卻是資料乾旱 (data drought)，因為我們缺乏跨越不同時空尺度的重複觀測和實驗資料來回答這些問題。為解決實際問題，必須將過往僅止於任意收集資料的方式轉變為解決問題收集資料的策略。要解決那些問題而收集特定的資料，可以參考生物多樣性公約 (Convention of Biological Diversity) 2020 年需達成之目標 (愛知目標 Aichi Targets)。講者認為有四類資料必須優先收集：1) 物種受威脅狀態製圖應用 (threat mapping application)，2) 保育優先 (conservation priorities)，3) 生物特徵資料 (trait-based data)，4) 生物多樣性公民科學家社群收集的資料 (iNaturalist data and services)。為了收集、分享、利用這些資料，也必須發展公開的生物多樣性資訊平台。

#### 6.1.2 政策－研究基礎架構介面：對生物多樣性資訊學的意義 (The policy - research infrastructure interface: implications for biodiversity informatics)

儘管我們對於物種、遺傳、生態、生物特徵等個別領域有許多的資料和了解，但是缺乏一個整合的知識體系協助我們了解、掌握這些知識之間的關聯，以解決全球生物多樣性快速消失的危機和挑戰。要解決全球性的難題挑戰，必須要全球的通力合作。現有分散的科學研究基礎設施必須進行整合，以提供更好的決策支援。對於如何建立一個全球合作的研究基礎設施架構，講者對大學、學術研究單位、政府、全球資料和基礎設施提供者、公眾參與社群提出了許多具體建議，並舉 GEOSS、GEO-BON 的架構為例。最後，強調全球公開、共享資料，盡可能的促進資料的使用，以解決全球生物多樣性消失的危機。

#### 6.1.3 我們半個生物圈 (不再) 不見了：瞭解生物多樣性、生物資訊學和生命本身的意義 (Half of Our Biosphere is (No Longer) Missing: Implications for Understanding Biodiversity, Bioinformatics, and Life Itself)

多樣性是生物學的必要條件，演化是生物學的核心，而基因是演化的基礎。過往，由於技術的限制，讓我們看到、理解生命或生物多樣性的現象是有所偏差、不完

整的。藉由最新基因定序技術而發展的基因組學 (genomics) 和元基因組學 (metagenomics) 改變了過往我們對於微生物多樣性的認知。過去，我們談的生物多樣性，往往忽略了微生物的重要性 (講者將微生物比喻微生物學的暗物質)。現在藉由元基因組學的研究，不僅可釐清微生物間演化的親緣關係，更讓我們深入瞭解微生物與其他生物間的密切關係。因此，21 世紀生物多樣性的研究必須涵蓋更多有關對微生物多樣性的瞭解。

## 6.2 分組討論與結論

研討會依據與會專家的專長、特質分為紅、綠、藍三組，進行三日不同議題的腦力激盪與討論。會議分組討論初步結果彙整的文件請參考：

<http://bit.ly/LLUhUb>

第一日的討論主軸是針對生物多樣性科學與資訊學的整合和資料管理如何增進我們對生物多樣性的瞭解、如何支援生物多樣性的基礎與應用研究進行討論。第二日的討論核心是以愛知目標 (Aichi Targets) 作為阻止全球生物多樣性消失議題的優先政策目標，針對生物多樣性資訊學和資料管理如何支援達成這些目標所需的資訊需求與全球行動。第三日的討論則為凝聚、建立解決這些重大議題的視野，針對前兩日平行分組討論關於生物多樣性資訊學的需求和機會，各團隊對於解決全球性生物多樣性消失危機所需的資料內容、資訊學和全球文化變革等各項需求，提出具體可行的建議做為未來 5-10 年內的 生物多樣性資訊學應用、發展的方向。



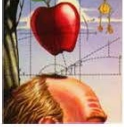



## 7 會議重點結論：

- 1) 應善加利用公眾的巨大潛能成為全球生物多樣性知識網的一部分，因為貢獻者和受益者都使用了最新的科技、社群網路和當地 (本土) 的知識。
- 2) 盡可能利用所有的科技以獲取關於物種間複雜交互作用關係的資料和知識。例如，捕食/被捕食者、寄生/宿主、傳粉者及其特徵。這些科技將包含聲音監測和各種遙測技術，並協助分析這些物種間交互作用的關係，以闡明這些交互作用所產生的生態系統服務的重要性的價值。
- 3) 大幅改進跨越不同時空尺度模式預測的能力，以評估在地球上任何地點特定環境因子變化生物多樣性的衝擊，及其對生態系服務和人類及生物群聚造成破壞的程度。

- 4) 利用連結開放資料 (linked open data, LOD) 技術將生物學名、博物館館藏紀錄、其他生物多樣性出現資料與 DNA、基因、微生物、甚至生態系監測的衛星遙測影像進行資料的連結，以拓展跨越不同尺度資料整合應用的能力。
- 5) 利用大量基因定序技術瞭解生存於空氣、河流、湖泊、海洋、土壤及其他生物體內數百萬、千萬種微生物的多樣性，及其在如何支持、控制地球的維生系統。

## 8 附錄一 主題演講簡報內容

### 8.1 Lucas Joppa: "Science-driven data"

<p>Microsoft <b>Research</b> Cambridge</p> <p>Computational Ecology and Environmental Science</p> <p>Lucas Joppa</p> <h3>Science-Driven Data</h3> <p>Biodiversity Informatics and Planetary Predictions</p> <p>The last time we are losing our confidence in the future, we are also beginning to close the door on our past? - President Jimmy Carter - "Crisis of Confidence" speech</p> <p>"This is where the vision needs to come in. This is what GBC should enable... to get the truly transformative ideas." - Donald Hobert's email</p>	<h3>Some Things We Can Predict With Confidence</h3> 										
<h3>We Do This Every Day</h3>  <p>Driving cars, flying in planes...</p>	<h3>Which Usually Works Out All Right</h3> <p>Chris Bishop didn't die because...</p>  <p>Data</p> <p>Theory</p> <p>Experiments</p> <p>combined to provide predictive certainty</p> 										
<h3>But We Also Do This Every Day</h3>  <p>Deforestation, ocean acidification, CO2 emissions, species extinctions</p>	<h3>And That Isn't A Good Idea</h3>  <p>To a large extent, we don't have any of that for the environment</p> <table border="1"> <thead> <tr> <th>Physics</th> <th>Ecology / Environment</th> </tr> </thead> <tbody> <tr> <td>Law of Inertia</td> <td>Species with small ranges are rare within them</td> </tr> <tr> <td>Law of Motion</td> <td>For X action, there will be Y reaction?</td> </tr> <tr> <td>Law of Gravity</td> <td>???</td> </tr> <tr> <td>Etc. Etc. Etc.</td> <td></td> </tr> </tbody> </table>	Physics	Ecology / Environment	Law of Inertia	Species with small ranges are rare within them	Law of Motion	For X action, there will be Y reaction?	Law of Gravity	???	Etc. Etc. Etc.	
Physics	Ecology / Environment										
Law of Inertia	Species with small ranges are rare within them										
Law of Motion	For X action, there will be Y reaction?										
Law of Gravity	???										
Etc. Etc. Etc.											

Still Unable To Answer *The* Key Questions About The Planet

- How many species are there?
- What do different species do?
- How fast are species going extinct?
- Which species matter?
- Will there be food and water for 10 billion people?
- Will ecosystems collapse?
- What will species do in the future?
- How will the earth system change?

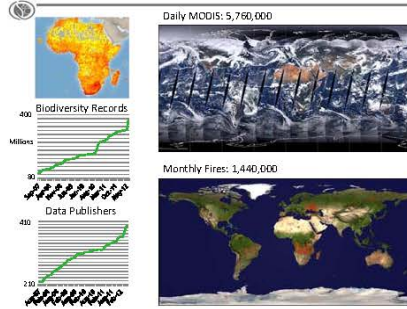
So Why Can't We Answer These Questions?



Is It A Data Problem? Yes and No...



Lots of Data!

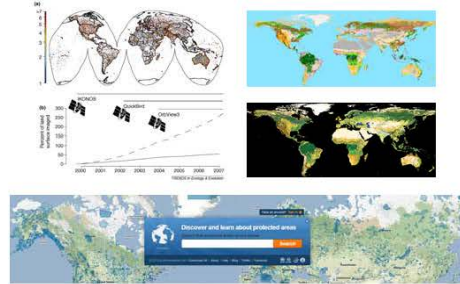


Is It A Data Problem?



Flying Blind...

We lack repeated measurements / observations / experiments across space and time





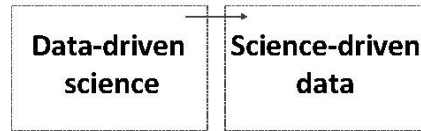


## The Big Problem

data

action

## Targeted & Incentivized Data Collection

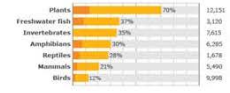


And by that I mean...

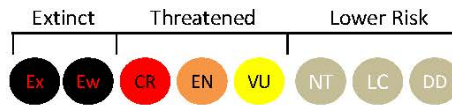
## Why We Are Here: CBD 2020

- Strategic Goal A:** Address the underlying causes of biodiversity loss by mainstreaming biodiversity across government and society  
 Target 4: By 2020, at the latest, Governments, business and stakeholders at all levels have taken steps to achieve or have implemented plans for sustainable production and consumption and have kept the impacts of use of natural resources well within safe ecological limits.
- Strategic Goal B:** Reduce the direct pressures on biodiversity and promote sustainable use  
 Target 5: By 2020, the rate of loss of all natural habitats, including forests, is at least halved and where feasible brought close to zero, and degradation and fragmentation is significantly reduced.
- Strategic Goal C:** Improve status of biodiversity by safeguarding ecosystems, species and genetic diversity  
 Target 11: By 2020, at least 17 per cent of terrestrial and inland water, and 10 per cent of coastal and marine areas, are conserved through effectively and equitably managed, ecologically representative and well-connected systems of Parks and other terrestrial and marine areas, and other effective area-based conservation measures  
 Target 12: By 2020 the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained.
- Strategic Goal D:** Enhance the benefits to all from biodiversity and ecosystem services  
 Target 15: By 2020, ecosystem resilience and the contribution of biodiversity to carbon stocks has been enhanced, through conservation and restoration, including restoration of at least 15 per cent of degraded ecosystems, thereby contributing to climate change mitigation and adaptation and to combating desertification.
- Strategic Goal E:** Implementation via planning, knowledge management and capacity building  
 Target 19: By 2020, knowledge, the science base and technologies relating to biodiversity, its values, functioning, status and trends, and the consequences of its loss, are improved, widely shared and transferred, and applied.

## Targeted Data I: IUCN Threat Mapping Application



- > Since 1994: Global standard, comprehensive, objective approach for evaluating the conservation status of plant and animal species.
- > Information required: 1) Spatial range & Population, 2) Threats, 3) Interventions.



## Threat Classifications

1. Residential and commercial development.
2. Agriculture and aquaculture.
3. Energy production and mining.
4. Transportation and service corridors.
5. Biological resource use
6. Human intrusions and disturbance.
7. Natural system modifications.
8. Invasive and other problematic species.
9. Pollution.
10. Geological events.
11. Climate change and severe weather.

Which of these data already exist?  
 -We asked that question.

Can we start gathering and sharing spatial information on threats and interventions?  
 -We built an application...

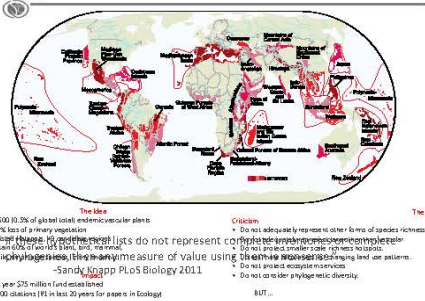
1. Residential and commercial development
  - 1.1. Housing and urban areas
    - 1.1.1. Urban sprawl
    - 1.1.2. Urban sprawl
    - 1.1.3. Urban sprawl
  - 1.2. Commercial and industrial areas
    - 1.2.1. Manufacturing sites, storage yards, open areas, other land uses
    - 1.2.2. Industrial sites
    - 1.2.3. Other commercial and industrial
2. Agriculture and Aquaculture
  - 2.1. Annual and perennial non-wood crops
    - 2.1.1. Shifting agriculture
    - 2.1.2. Sustainable farming
    - 2.1.3. Agriculture farming
    - 2.1.4. Agriculture farming
    - 2.1.5. Agriculture farming
  - 2.2. Wood and pulp plantations
  - 2.3. Livestock, farming and ranching
    - 2.3.1. Hunting
    - 2.3.2. Sustainable agriculture
    - 2.3.3. Agriculture farming
    - 2.3.4. Agriculture farming
    - 2.3.5. Agriculture farming
  - 2.4. Marine and freshwater aquaculture
    - 2.4.1. Sustainable agriculture
    - 2.4.2. Sustainable agriculture
    - 2.4.3. Sustainable agriculture

## Targeted Data I: Threat Mapping Application

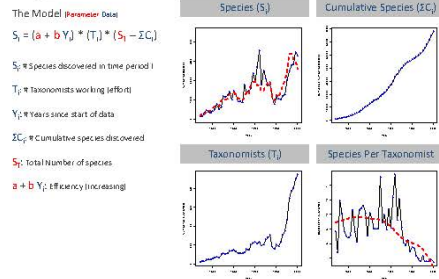




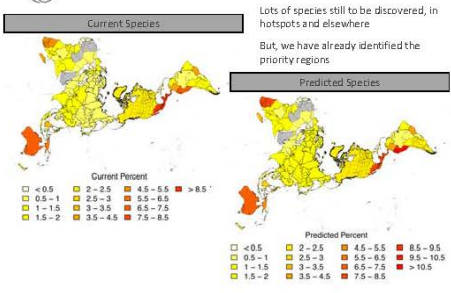
### Targeted Data II: Conservation Priorities



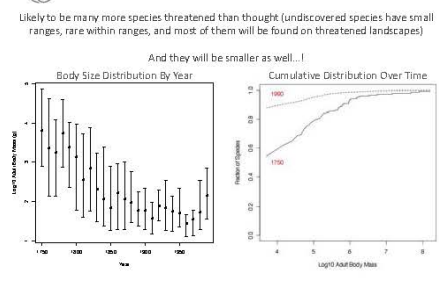
### Targeted Data II: Conservation Priorities



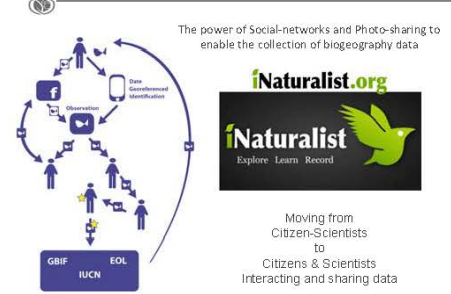
### Targeted Data II: Conservation Priorities



### Targeted Data III: Trait-Based Data



### Targeted Data IV: iNaturalist.org



### Targeted Data IV: iNaturalist.org services

**Expert Assistance**

**Checklists**

And checklists are great for data targeting!

### Biodiversity Informatics Platforms

### Computational Ecology @ Microsoft Research

CEES aims to develop the new concepts, methods, and software tools that are needed to produce *predictive models of environmental systems*

**Computational Ecology and Environmental Sciences**

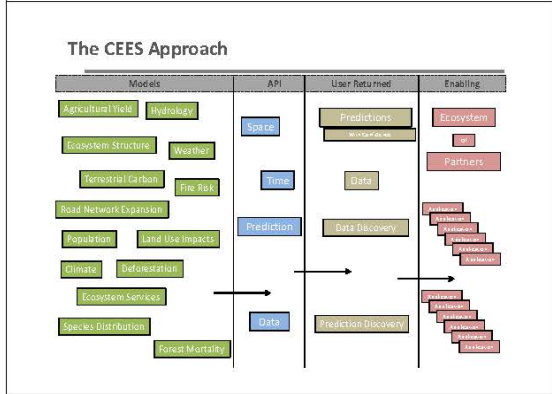
Win: Just might help Microsoft evaluate how a new environmental prediction service

Yousef: Will still have advanced important areas of environmental science

Enabling Technologies: Agricultural Yields, Road Network Expansion, Ecological Networks, Animal Movements, Balanced Co-empirical Approaches, Species Distributions, Forest Mortality, Environmental Trends, Species Extinctions, Ecosystem Modeling, Carbon Cycle

Yousef: Will still have created useful tools for science and for Microsoft

Flascon: HYLL, AMP Lact, Network3D, Dynamic Dataset Viewer, Project Swertery



### Fetch! (Climate, Carbon)...

## 8.2 Daan du Toit: "The policy - research infrastructure interface: implications for biodiversity informatics"

 <p><b>Research Infrastructure – Policy Interface: Implications for Biodiversity Informatics</b></p> <p>Daan du Toit Senior S&amp;T Representative to the EU</p> <p>Presentation to GBIC2012</p> <p>Department of Science and Technology REPUBLIC OF SOUTH AFRICA</p>	 <p><b>Scope</b></p> <ul style="list-style-type: none"> <li>• <b>GBIC: Coordinate action to mobilize required data and informatics capacities to support biodiversity science and address policy objectives</b> <ul style="list-style-type: none"> <li>– How best assemble and interpret data for Aichi targets – i.e. secure optimal research infrastructure support</li> </ul> </li> <li>• <b>Presentation: Interface research infrastructure (RI) and policy</b> <ul style="list-style-type: none"> <li>– “Enabling policy environment” for RI</li> <li>– Policy-maker “expectations” from RI</li> </ul> </li> </ul>
 <p><b>RI Definition</b></p> <ul style="list-style-type: none"> <li>• <b>Single-sited, distributed or virtual facilities, resources or services to support knowledge generation, solve challenges, innovation</b> <ul style="list-style-type: none"> <li>– major equipment or sets of instruments</li> <li>– knowledge resources: collections, archives, databanks</li> <li>– e-Infrastructures: <ul style="list-style-type: none"> <li>• structured information systems for data management, enabling information and communication</li> <li>• technology-based Infrastructures: network, grid, computing, software and middleware</li> </ul> </li> </ul> </li> </ul>	 <p><b>STI Policy Context for RI</b></p> <ul style="list-style-type: none"> <li>• Focus on grand / global challenges <ul style="list-style-type: none"> <li>– Climate change, energy security, pandemic disease, etc.</li> </ul> </li> <li>• Science knows no borders <ul style="list-style-type: none"> <li>– Increased international cooperation</li> <li>– “Networked science”</li> <li>– Joint programming</li> <li>– Multi-disciplinary</li> <li>– Public-private</li> </ul> </li> <li>• Data as service: riding the tide / open access</li> <li>• Bridging the innovation chasm / knowledge triangle</li> <li>• Funding: Opportunity – science drives growth and development, but also constraint – competing priorities, look non-conventional sources</li> </ul>
 <p><b>Overview presentation</b></p> <ul style="list-style-type: none"> <li>• RI policy developments relevant for biodiversity informatics initiatives <ul style="list-style-type: none"> <li>– Specific focus data (Group of Senior Officials)</li> </ul> </li> <li>• (Broader) policy expectations from RI <ul style="list-style-type: none"> <li>– Including non-scientific specific impacts</li> </ul> </li> <li>• Reflection on GEOSS</li> <li>• Concluding thoughts</li> <li>• <i>Informed by experience:</i> <ul style="list-style-type: none"> <li>– OECD Global Science Forum</li> <li>– International Conference on Research Infrastructures</li> <li>– G8+5 Group of Senior Officials on Global Research Infrastructures</li> <li>– Development of Global Earth Observation System of Systems (GEOSS)</li> </ul> </li> </ul>	 <p><b>International RI investment / cooperation framework</b></p> <ul style="list-style-type: none"> <li>• RI roadmaps results of strategic, long-term, policy-relevant planning exercise <ul style="list-style-type: none"> <li>– E.g. ESFRI</li> </ul> </li> <li>• Criteria for determining RI priorities include: <ul style="list-style-type: none"> <li>– Benefits to as many Grand Challenges as possible</li> <li>– Those under threat requiring immediate attention</li> <li>– Potential to leverage international partnership, including comparative advantages of partners, e.g.: <ul style="list-style-type: none"> <li>• Knowledge</li> <li>• Geographic</li> <li>• Resource</li> </ul> </li> </ul> </li> </ul>

<p><b>Group of Senior Officials on Global Research Infrastructures</b></p> <ul style="list-style-type: none"> <li>• Framework for a coherent and coordinated world-wide development and operation of global research infrastructures</li> <li>• To inform decisions on (based best practice): <ul style="list-style-type: none"> <li>– prioritisation</li> <li>– efficient governance structures</li> <li>– appropriate funding schemes</li> <li>– policies for access and utilisation</li> <li>– Etc.</li> </ul> </li> </ul> 	<p><b>GSO Draft Recommendations</b></p> <ul style="list-style-type: none"> <li>• Core purpose of global research infrastructures</li> <li>• Defining project partnerships for effective management</li> <li>• Defining scope, schedule and cost</li> <li>• Project management</li> <li>• Funding management</li> <li>• Periodic reviews</li> <li>• Access based on merit review</li> <li>• E-infrastructure</li> <li>• Data exchange</li> <li>• Clustering of research infrastructures</li> <li>• International mobility</li> <li>• Technology transfer and intellectual property</li> <li>• Monitoring socio-economic impact</li> </ul> 
<p><b>GSO: e-Infrastructures / Data exchange</b></p> <ul style="list-style-type: none"> <li>• Special working group on data</li> <li>• Promote integrated use of advanced e-infrastructures and services for: <ul style="list-style-type: none"> <li>– accessing and processing, and curating data</li> <li>– remote participation (interaction) and access to scientific experiments</li> </ul> </li> <li>• Encourage global scientific data infrastructure providers and users to promote <ul style="list-style-type: none"> <li>– data exchange and interoperability of data across disciplines and national boundaries to broaden the scientific reach of individual data sets</li> </ul> </li> </ul> 	<p><b>GSO Working Group on Data</b></p> <ul style="list-style-type: none"> <li>• Call for action: <ul style="list-style-type: none"> <li>- Universities</li> <li>- Global scientific data infrastructure providers and users</li> <li>- Research institutions and funding entities</li> <li>- Communities of as well as individual scientists and researchers</li> <li>- Governments</li> <li>- Global governance frameworks</li> <li>- Libraries and publishers</li> <li>- Networks of repositories, libraries and data-centres</li> <li>- Public and private sector developers of advanced applications, tools and services</li> </ul> </li> <li>• Does not replace or supersede national or regional policies</li> </ul> 
<p><b>Summary GSO Data WG draft recommendations</b></p> <ul style="list-style-type: none"> <li>• Global scientific data infrastructure providers, research institutions and funding entities (should): <ul style="list-style-type: none"> <li>– address all components of data life cycle: <ul style="list-style-type: none"> <li>• acquisition,</li> <li>• curation,</li> <li>• metadata,</li> <li>• provenance,</li> <li>• persistent identifiers,</li> <li>• authorization,</li> <li>• authentication and data integrity: including development of approaches that provide a common look and feel to data discovery across disciplines and</li> </ul> </li> <li>– reduce learning curve required to achieve productivity</li> </ul> </li> </ul> 	<p><b>Draft recommendations: Training</b></p> <p><b>I. Universities and research organisations:</b></p> <ul style="list-style-type: none"> <li>• <b>actively promote stewardship</b> in sharing and re-using data and information <b>by making it a factor in scientific career advancement.</b></li> </ul> <p><b>II. Global scientific data infrastructure providers, research institutions and funding entities:</b></p> <ul style="list-style-type: none"> <li>• <b>support development &amp; training of new cohorts of data-intensive computational science expertise</b> (including data specialists, technicians and data managers).</li> </ul> 



### Draft recommendations: User communities

III. **Communities** in different fields of science and humanities, research facilities producing data:

- **take active role in defining concrete short and long term requirements for the underlying scientific data infrastructures for data**

IV. **Scientists and researchers:**

- **embrace open science** implementing and realising a scientific culture of fair sharing and respect of other's work ("good scientific practice").



### Draft recommendations: Infrastructure and Access

V. **Governments:**

- **provide adequate funding** to develop interconnected seamless global data infrastructure(s) underpinning scientific culture of open and fair exchange—supporting scientists to exchange data, publications and knowledge without barriers

VI. **Global governance frameworks:**

- **eliminate unnecessary barriers** to accessing data and promote recognition and reputation mechanisms which encourage it.



### Draft recommendations: Infrastructure and Access (2)

V. **Libraries and Publishers:**

- **ensure that data are readily discoverable** by all potential users **without requiring specialist expertise** and prior knowledge of existence of the data sources. Key: standardisation of metadata for cataloguing of data itself and for the means of aggregation, collection and indexing these standards providing a common look and feel to data discovery across disciplines



### Draft recommendations: Interoperability

VIII. **Global scientific data infrastructure providers and users:**

- **establish international forum for data interoperability to facilitate exchange and interoperability of data** across disciplines and national boundaries by producing high quality, relevant technical documents and procedures that influence the way researchers store, use, and manage data.
- **Forum directed by adequate international governance board** to set strategies and priorities:
  - > **managed by users who focus on operational issues** associated with the exchange and transmission of data. Adopt process similar to IETF, focus on "rough and consensus and running code" to insure that data sharing takes place from outset rather than waiting for standards to be finalized and approved.



### Draft recommendations: Interoperability (2)

IX. **Networks of repositories, libraries and data-centres:**

- **interoperate at global level** with high levels of dependability and trust, guided by international standards;
- **profit from and optimise synergies with network and computing infrastructures**, and
- **engage user communities of researchers in defining useful standards, services and platforms** to be developed and maintained.



### Draft recommendations: Development of new tools


X. **Public and Private sector developers of advanced applications, tools and services:**

- **have incentive** to innovate.
- In future—even more important as increasingly complex data will not be understandable without access to and use of yet-to-be-developed analysis tools, visualizations, decision-making support, models, simulations, etc...

XI. **Global scientific data infrastructure providers, research institutions and national funding agencies:**

- **encourage leveraging and building on already existing resources to be economically efficient.**



 <p><b>Broader “expectations” from RI</b></p> <ul style="list-style-type: none"> <li>• Enhance scientific tools to respond to global challenges</li> <li>• Leverage socio-economic benefits, including driver for innovation in key enabling technologies</li> <li>• Instrument for encouraging inclusive international cooperation</li> <li>• Promote science diplomacy</li> <li>• Vehicle for efficient joint investments from multiple funding instruments</li> <li>• Enable equitable brain circulation, facilitate retention of skill</li> <li>• Support human capital development for research and technical training</li> <li>• Promote public understanding and awareness of science, specifically for youth</li> </ul>	<p>GEOSS an example of an emerging global public good infrastructure</p> <ul style="list-style-type: none"> <li>➤ <b>Coordinate and Sustain Observation Systems</b></li> <li>➤ <b>Provide Easier &amp; More Open Data Access</b></li> <li>➤ <b>Foster Use through Science, Applications and Capacity Building</b></li> </ul>  <p>... to answer Society's need for informed decision making</p>
<p><b>GEO Bon Example of in-situ infrastructure</b></p> 	<p><b>GEOSS Data Sharing Principles</b></p> <ul style="list-style-type: none"> <li>• The GEOSS 10-Year Implementation Plan (2005) states:</li> <li>• <b>“The societal benefits of Earth observations cannot be achieved without data sharing.”</b></li> </ul> <p>There will be full and open exchange of data, metadata and products shared within GEOSS, recognizing relevant international instruments and national policies and legislation.</p> <p>All shared data, metadata and products will be made available with minimum time delay and at minimum cost.</p> <p>All shared data, metadata and products being free of charge or no more than cost of reproduction will be encouraged for research and education.</p>
<p><b>GEOSS Dissemination of Information</b></p> <ul style="list-style-type: none"> <li>• GEOSS will disseminate information and analyses directly to users</li> <li>• GEO has developed the GEOPortal as a single Internet gateway</li> <li>• The purpose of GEOPortal is to make it easier to integrate diverse data sets <ul style="list-style-type: none"> <li>– identify relevant data and portals of contributing systems</li> <li>– access models and other decision-support tools</li> </ul> </li> </ul>	<p><b>GEOSS harnesses optimal support from e-Infrastructures</b></p> <ul style="list-style-type: none"> <li>• Interlinking observation systems requires common standards for architecture and data sharing</li> <li>• Related networking technologies are critical for meeting the requirements of GEOSS</li> <li>• Satellite communications, ad hoc and sensor communications and networking, free-space optical (FSO) communications systems, large data storages, high-performance computing and broadband wireless networks</li> </ul>

## 8.3 Robert Robbins: "Half of Our Biosphere is (No Longer) Missing: Implications for Understanding Biodiversity, Bioinformatics, and Life Itself"

<p> <small>FOUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL BIOINFORMATIC COMMUNITY</small> </p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: center;"><small>PROVIDING DEVELOPMENT, INTERACTION, AND ASSISTANCE TO GENOMICS AND BIOINFORMATIC DATA STANDARDS</small></p> <h1 style="text-align: center;">Half of Our Biosphere Is Missing</h1> <p style="text-align: center;"><small>ORBI08C - 2 JUL 2012 - Copenhagen, ©2012, R.J. Robbins</small></p> <p style="text-align: right;"><small>1</small></p>	<p> <small>FOUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL BIOINFORMATIC COMMUNITY</small> </p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: center;"><small>PROVIDING DEVELOPMENT, INTERACTION, AND ASSISTANCE TO GENOMICS AND BIOINFORMATIC DATA STANDARDS</small></p> <h1 style="text-align: center;">Half of Our Biosphere Is Missing</h1> <p style="text-align: center;"><i>No Longer</i></p> <h2 style="text-align: center;">Implications for Understanding Biodiversity, Bioinformatics, and Life Itself</h2> <p style="text-align: center;"><small>ORBI08C - 2 JUL 2012 - Copenhagen, ©2012, R.J. Robbins</small></p> <p style="text-align: right;"><small>2</small></p>
<p> <small>FOUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL BIOINFORMATIC COMMUNITY</small> </p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: center;"><small>PROVIDING DEVELOPMENT, INTERACTION, AND ASSISTANCE TO GENOMICS AND BIOINFORMATIC DATA STANDARDS</small></p> <h1 style="text-align: center;">Half of Our Biosphere Is Missing</h1> <p style="text-align: center;"><i>No Longer</i></p> <h2 style="text-align: center;">Implications for Understanding Biodiversity, Bioinformatics, and Life Itself</h2> <p style="text-align: center;">Robert J. Robbins RJR8222@gmail.com</p> <p style="text-align: center;"> <a href="http://www.rj-robbins.com/slides/RJR-GBIC-2012.pdf">http://www.rj-robbins.com/slides/RJR-GBIC-2012.pdf</a> </p> <p style="text-align: center;"><small>ORBI08C - 2 JUL 2012 - Copenhagen, ©2012, R.J. Robbins</small></p> <p style="text-align: right;"><small>3</small></p>	<p> <small>FOUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL BIOINFORMATIC COMMUNITY</small> </p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: center;"><small>PROVIDING DEVELOPMENT, INTERACTION, AND ASSISTANCE TO GENOMICS AND BIOINFORMATIC DATA STANDARDS</small></p> <h3>Abstract:</h3> <p>In the last few years, improving genomic and metagenomic tools have been revealing details about the previously invisible microbial world. Although these new findings are yielding very important insights about global biodiversity, some may be difficult to accommodate in current biological models. Half of the world's biomass and by far the majority of its biodiversity are, for the first time, becoming available for study. Every month, startling results appear in the literature. Not only do free-living microbes represent a distinctly alien way of life, commensal microbes are proving to have profound effects on their host organisms, affecting things ranging from mate choice in <i>Drosophila</i>, to pain tolerance in mice, to niche partitioning in ants. The emerging pervasive influence of commensal microbes suggests that to fully understand the biology of any organism, we must take into account its interactions with its associated microbiota. It seems, we are all lichens now. The resulting conceptual adjustments will offer great opportunities for expanding our understanding of the biosphere, but will offer real challenges to our current view of biodiversity and will greatly complicate the informatics tools needed to document biodiversity. Not only will biodiversity informatics projects need to deal with an explosion in the amount of biodiversity-relevant data, they may well need to accommodate data that are of a conceptually different form. As any informatics professional knows, making changes to an underlying data model is always difficult and fraught with risk. Making changes to conceptual base classes is the hardest of all. Welcome to the world of 21st Century biodiversity.</p> <p style="text-align: center;"><small>ORBI08C - 2 JUL 2012 - Copenhagen, ©2012, R.J. Robbins</small></p> <p style="text-align: right;"><small>4</small></p>
<p> <small>FOUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL BIOINFORMATIC COMMUNITY</small> </p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: center;"><small>PROVIDING DEVELOPMENT, INTERACTION, AND ASSISTANCE TO GENOMICS AND BIOINFORMATIC DATA STANDARDS</small></p> <h3>GBIF and Biodiversity:</h3> <ul style="list-style-type: none"> <li>• What is "biodiversity" and how should biodiversity information be managed?</li> <li>• Sequencing is getting better and faster at an incredible rate. What is the relevance to biodiversity studies?</li> <li>• Science is a "light's better" endeavor. When the light changes, the science changes.</li> <li>• The light is changing: Biological dark matter is becoming visible.</li> <li>• Reality is not negotiable I: Examples from genetics &amp; genomics.</li> <li>• Reality is not negotiable II: The future of biodiversity.</li> <li>• Welcome to the world of 21st Century biodiversity.</li> </ul> <p style="text-align: center;"><small>ORBI08C - 2 JUL 2012 - Copenhagen, ©2012, R.J. Robbins</small></p> <p style="text-align: right;"><small>5</small></p>	<p> <small>FOUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL BIOINFORMATIC COMMUNITY</small> </p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: center;"><small>PROVIDING DEVELOPMENT, INTERACTION, AND ASSISTANCE TO GENOMICS AND BIOINFORMATIC DATA STANDARDS</small></p> <h3>What is "biodiversity" and how should biodiversity information be managed?</h3> <p style="text-align: center;"><small>ORBI08C - 2 JUL 2012 - Copenhagen, ©2012, R.J. Robbins</small></p> <p style="text-align: right;"><small>6</small></p>

What is “biodiversity” and how should biodiversity information be managed?

*Addressing this question is, essentially, the purpose of this meeting.*



08F080C - 24 JUL 2012 - 09:45am, #2012, R.J. Robins

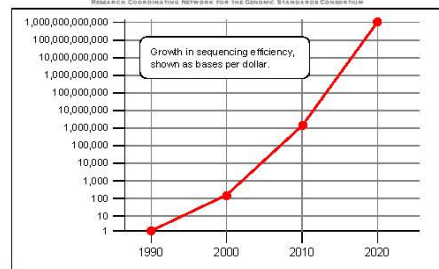
7

Sequencing Improvement  
Is Astounding !!



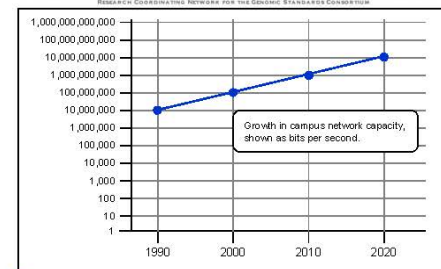
08F080C - 24 JUL 2012 - 09:45am, #2012, R.J. Robins

8



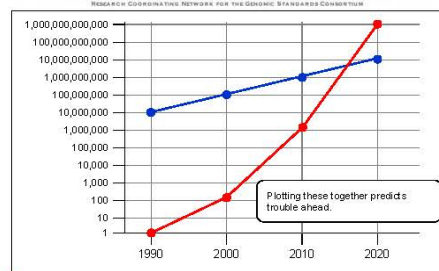
08F080C - 24 JUL 2012 - 09:45am, #2012, R.J. Robins

9



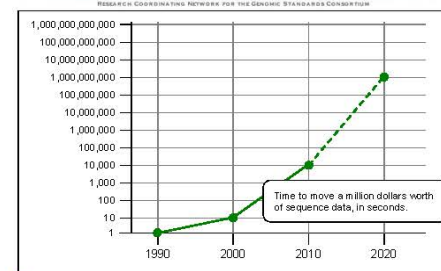
08F080C - 24 JUL 2012 - 09:45am, #2012, R.J. Robins

10



08F080C - 24 JUL 2012 - 09:45am, #2012, R.J. Robins

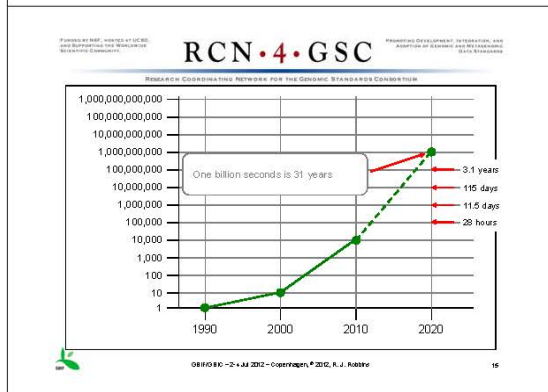
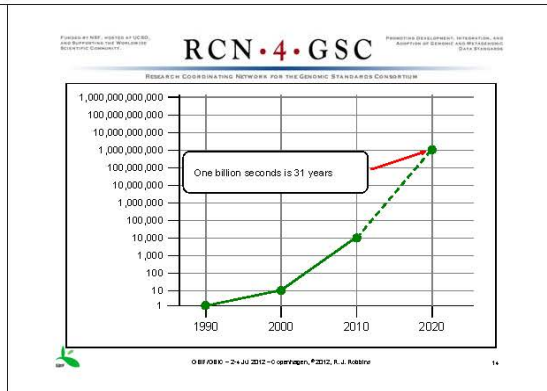
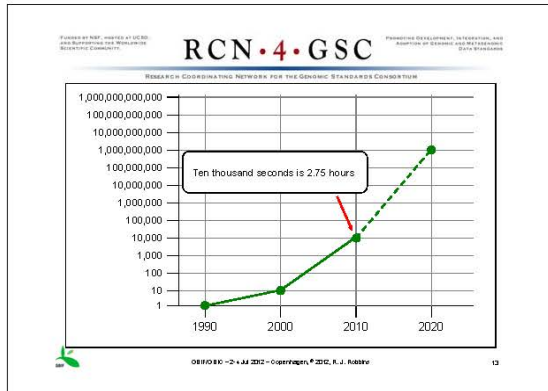
11



08F080C - 24 JUL 2012 - 09:45am, #2012, R.J. Robins

12





FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Bottom Line:

- There will be a lot more sequence data in the future than there is now.

0810080 - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble
16

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Bottom Line:

- There will be a lot more sequence data in the future than there is now.
- Incorporating generic sequence data into biodiversity informatics will be technically and logistically challenging.

0810080 - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble
17







FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS







RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Bottom Line:

- There will be a lot more sequence data in the future than there is now.
- Incorporating generic sequence data into biodiversity informatics will be technically and logistically challenging.
- Incorporating biodiversity-specific sequence data into biodiversity informatics may also be conceptually and logically challenging.

0810080 - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble
18

<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COORDINATING NETWORK</small>  <b>RCN•4•GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <h2 style="text-align: center;">Sequencing and Biodiversity</h2> <h3 style="text-align: center;">I. The Basics</h3> <p style="text-align: center;">  <small>08F08K - 24 JUL 2012 - Copenhagen, #2012, R.J. Robble</small> <span style="float: right;">19</span> </p>	<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COORDINATING NETWORK</small>  <b>RCN•4•GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <h3>Why (meta)genomics and biodiversity?</h3> <ul style="list-style-type: none"> <li>• Biodiversity is less a field of biology than a perspective (that of variance) into biology.</li> <li>• Diversity is a sine qua non of biology; no diversity, no evolution.</li> <li>• Genetics / genomics are equally central to biology – genetics is the study of the hereditary machinery, the basis of heritable variation, the raw material for evolution, the ultimate source of biodiversity.</li> </ul> <p style="text-align: center;">  <small>08F08K - 24 JUL 2012 - Copenhagen, #2012, R.J. Robble</small> <span style="float: right;">20</span> </p>
<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COORDINATING NETWORK</small>  <b>RCN•4•GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <h3>Why (meta)genomics and biodiversity?</h3> <ul style="list-style-type: none"> <li>• Biodiversity is less a field of biology than a perspective</li> </ul> <div style="border: 1px solid black; border-radius: 15px; padding: 10px; width: fit-content; margin: 10px auto;"> <p style="text-align: center;">The connection between genomics/ metagenomics and biodiversity seems obvious and profound.</p> </div> <p style="text-align: center;">  <small>08F08K - 24 JUL 2012 - Copenhagen, #2012, R.J. Robble</small> <span style="float: right;">21</span> </p>	<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COORDINATING NETWORK</small>  <b>RCN•4•GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <h3>The Basics:</h3> <ul style="list-style-type: none"> <li>• As sequencing gets cheaper, its practical applicability to biodiversity will increase.</li> </ul> <p style="text-align: center;">  <small>08F08K - 24 JUL 2012 - Copenhagen, #2012, R.J. Robble</small> <span style="float: right;">22</span> </p>
<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COORDINATING NETWORK</small>  <b>RCN•4•GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <h3>The Basics:</h3> <ul style="list-style-type: none"> <li>• As sequencing gets cheaper, its practical applicability to biodiversity will increase.</li> <li>• Barcode-type data are useful in a diversity-diagnostic sense.</li> </ul> <p style="text-align: center;">  <small>08F08K - 24 JUL 2012 - Copenhagen, #2012, R.J. Robble</small> <span style="float: right;">23</span> </p>	<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COORDINATING NETWORK</small>  <b>RCN•4•GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <h3>The Basics:</h3> <ul style="list-style-type: none"> <li>• As sequencing gets cheaper, its practical applicability to biodiversity will increase.</li> <li>• Barcode-type data are useful in a diversity-diagnostic sense.</li> <li>• Metagenomic tools allow a broad diversity assessment in a single test.</li> </ul> <p style="text-align: center;">  <small>08F08K - 24 JUL 2012 - Copenhagen, #2012, R.J. Robble</small> <span style="float: right;">24</span> </p>

<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small> </p> <p style="text-align: center;"> <b>RCN•4•GSC</b> </p> <p style="text-align: center;"> <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>The Basics:</b></p> <ul style="list-style-type: none"> <li>• As sequencing gets cheaper, its practical applicability to biodiversity will increase.</li> <li>• Barcode-type data are useful in a diversity-diagnostic sense.</li> <li>• Metagenomic tools allow a broad diversity assessment in a single test.</li> <li>• Traditional biodiversity thinking will need to be extended to include sequence data; e.g., will we need a type sequence?</li> </ul> <p>  <small>08FJ08C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble</small> </p>	<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small> </p> <p style="text-align: center;"> <b>RCN•4•GSC</b> </p> <p style="text-align: center;"> <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>The Basics:</b></p> <ul style="list-style-type: none"> <li>• <b>Geo-referenced genetic data can provide evidence of patterns of origin and distribution of new genes.</b></li> </ul> <p>  <small>08FJ08C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble</small> </p>
<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small> </p> <p style="text-align: center;"> <b>RCN•4•GSC</b> </p> <p style="text-align: center;"> <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>The Basics:</b></p> <ul style="list-style-type: none"> <li>• Geo-referenced genetic data can provide evidence of patterns of origin and distribution of new genes.</li> <li>• Should species-abundance maps be extended to include geno-clines?</li> </ul> <p>  <small>08FJ08C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble</small> </p>	<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small> </p> <p style="text-align: center;"> <b>RCN•4•GSC</b> </p> <p style="text-align: center;"> <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>The Basics:</b></p> <ul style="list-style-type: none"> <li>• Geo-referenced genetic data can provide evidence of patterns of origin and distribution of new genes.</li> <li>• Should species-abundance maps be extended to include geno-clines?</li> <li>• and so on...</li> </ul> <p>  <small>08FJ08C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble</small> </p>
<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small> </p> <p style="text-align: center;"> <b>RCN•4•GSC</b> </p> <p style="text-align: center;"> <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <b>Sequencing and Biodiversity</b>  <b>II. New Insights</b> </p> <p>  <small>08FJ08C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble</small> </p>	<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small> </p> <p style="text-align: center;"> <b>RCN•4•GSC</b> </p> <p style="text-align: center;"> <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>New Insights:</b></p> <ul style="list-style-type: none"> <li>• Sequence analysis was responsible for the most important biodiversity discovery of the last hundred years.</li> </ul> <p>  <small>08FJ08C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robble</small> </p>

RCN·4·GSC

Research Coordinating Network for the Genomic Standards Consortium

**New Insights:**

- Sequence analysis was responsible for the most important biodiversity discovery of the last hundred years.
- Newly emerging sequence analysis tools will allow us to study vast swaths of biodiversity that have previously been invisible.

© 2012 R. J. Robbins

RCN·4·GSC

Research Coordinating Network for the Genomic Standards Consortium

**New Insights:**

- Sequence analysis was responsible for the most important biodiversity discovery of the last hundred years.
- Newly emerging sequence analysis tools will allow us to study vast swaths of biodiversity that have previously been invisible.
- Findings from genomic and metagenomic studies of biodiversity may force some MAJOR reassessments of basic biological concepts.

© 2012 R. J. Robbins

RCN·4·GSC

Research Coordinating Network for the Genomic Standards Consortium

© 2012 R. J. Robbins

RCN·4·GSC

Research Coordinating Network for the Genomic Standards Consortium

**Major Biodiversity Discovery !**

© 2012 R. J. Robbins

RCN·4·GSC

Research Coordinating Network for the Genomic Standards Consortium

**Major Biodiversity Discovery !**

Not really.  
It's just another frog ...

© 2012 R. J. Robbins

RCN·4·GSC

Research Coordinating Network for the Genomic Standards Consortium

**Major Biodiversity Discovery !**

**Phylogenetic structure of the prokaryotic domain: The primary kingdoms**  
(archaeobacteria/eubacteria/eukaryotes/16S ribosomal RNA/molecular phylogeny)

CARL R. WOESE AND GEORGE E. FOX\*

Department of Genetics and Evolutionary, University of Illinois, Urbana, Illinois 61801

Communicated by T. M. Sonneborn, August 18, 1977

**ABSTRACT** A phylogenetic analysis based upon ribosomal RNA sequence characterization reveals that living organisms are of three theoretical levels of descent: (1) the eubacteria, comprising all typical members of the eubacteria, bacteria, cyanobacteria, and all of the ultrabacteria, now considered prokaryotic members of the eubacteria; (2) the ultrabacteria, comprising all members of the ultrabacteria, now considered members of the eubacteria; and (3) the eukaryotes, comprising all members of the eukaryotes, now considered members of the eubacteria.

The biologist has customarily structured his world in terms of certain basic dichotomies. Central to these is the distinction between plants and animals, which initially had been considered plants, revealed both plants and animals had descended from a common ancestor. The discovery that plants and animals descended from a common ancestor led to a yet more basic dichotomy, that of plants and animals, revealed both plants and animals had descended from a common ancestor.

to construct phylogenetic classifications between domains. Prokaryotic kingdoms are not comparable to eukaryotic ones. This should be recognized by an appropriate terminology. The highest phylogenetic unit in the prokaryotic domain we think should be called an "ultrakingdom"—an eubacteria "primary kingdom." This would recognize the significant distinctions between prokaryotic and eukaryotic kingdoms and emphasize that the former have greater evolutionary status.

The passage from one domain to a higher one has become a central problem. Initially one would like to know whether this is a frequent or a rare unique evolutionary event. It is traditionally assumed—without evidence—that the eukaryotic domain has arisen but once, all eukaryotes stem from a common ancestor.

© 2012 R. J. Robbins

RCN·4·GSC

Major Biodiversity Discovery!

Phylogenetic structure of the prokaryotic domain: The primary kingdoms

Now this was a very big deal...

08F08C - 24 Jul 2012 - 0 paper, #2012, R.J. Robble

RCN·4·GSC

LOGIC:

One could measure the similarity between two text documents by chopping them up into, say, ten-word phrases, and then asking what percentage of ten-word phrases were present in common between two documents.

If every phrase occurred in both documents, the score would be 1.0, if no phrases occurred in both documents, the score would be 0.0.

08F08C - 24 Jul 2012 - 0 paper, #2012, R.J. Robble

RCN·4·GSC

Evolution: Wase and Fox

Table 1. Association coefficients ( $S_{ij}$ ) between representative members of the three primary kingdoms

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. <i>Saccharomyces cerevisiae</i> , 18S	—	0.29	0.33	0.05	0.06	0.09	0.11	0.08	0.11	0.11	0.08	0.08	0.08
2. <i>Zinnia mexicana</i> , 18S	0.29	—	0.36	0.10	0.09	0.10	0.09	0.11	0.10	0.10	0.12	0.07	0.07
3. <i>L. cell</i> , 18S	0.33	0.36	—	0.06	0.06	0.07	0.07	0.09	0.06	0.10	0.10	0.09	0.07
4. <i>Escherichia coli</i>	0.05	0.10	0.06	—	0.24	0.25	0.28	0.26	0.21	0.11	0.12	0.07	0.12
5. <i>Chlorella vibrioforme</i>	0.06	0.06	0.06	0.24	—	0.22	0.22	0.19	0.19	0.06	0.07	0.06	0.09
6. <i>Bacillus firmus</i>	0.08	0.06	0.07	0.25	0.22	—	0.24	0.26	0.20	0.11	0.13	0.06	0.12
7. <i>Corynebacterium diptheriae</i>	0.09	0.10	0.07	0.26	0.23	0.24	—	0.21	0.21	0.12	0.12	0.09	0.10
8. <i>Aphanogony 8714</i>	0.11	0.09	0.09	0.26	0.23	0.23	0.21	—	0.31	0.11	0.11	0.10	0.10
9. <i>Chlamydia (Genus)</i>	0.08	0.11	0.06	0.21	0.19	0.21	0.21	0.31	—	0.14	0.12	0.10	0.12
10. <i>Methanobacterium thermoautotrophicum</i>	0.11	0.10	0.10	0.11	0.06	0.11	0.12	0.11	0.14	—	0.51	0.25	0.20
11. <i>M. thermoautotrophicum strain M3</i>	0.11	0.10	0.10	0.12	0.07	0.12	0.12	0.12	0.12	0.51	—	0.25	0.24
12. <i>Methanobacterium sp. Carotococcus JR-1</i>	0.08	0.12	0.09	0.07	0.06	0.06	0.09	0.10	0.10	0.25	0.25	—	0.32
13. <i>Methanopyrus kandleri</i>	0.08	0.07	0.07	0.12	0.08	0.12	0.10	0.12	0.20	0.24	0.32	0.32	—

RESULTS

08F08C - 24 Jul 2012 - 0 paper, #2012, R.J. Robble

RCN·4·GSC

Range: 0.29-0.36

Range: 0.05-0.13

Table 1. Association coefficients ( $S_{ij}$ ) between representative members of the three primary kingdoms

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. <i>Saccharomyces cerevisiae</i> , 18S	—	0.29	0.33	0.05	0.06	0.09	0.11	0.08	0.11	0.11	0.08	0.08	0.08
2. <i>Zinnia mexicana</i> , 18S	0.29	—	0.36	0.10	0.09	0.10	0.09	0.11	0.10	0.10	0.12	0.07	0.07
3. <i>L. cell</i> , 18S	0.33	0.36	—	0.06	0.06	0.07	0.07	0.09	0.06	0.10	0.10	0.09	0.07
4. <i>Escherichia coli</i>	0.05	0.10	0.06	—	0.24	0.25	0.28	0.26	0.21	0.11	0.12	0.07	0.12
5. <i>Chlorella vibrioforme</i>	0.06	0.06	0.06	0.24	—	0.22	0.22	0.19	0.19	0.06	0.07	0.06	0.09
6. <i>Bacillus firmus</i>	0.08	0.06	0.07	0.25	0.22	—	0.24	0.26	0.20	0.11	0.13	0.06	0.12
7. <i>Corynebacterium diptheriae</i>	0.09	0.10	0.07	0.26	0.23	0.24	—	0.21	0.21	0.12	0.12	0.09	0.10
8. <i>Aphanogony 8714</i>	0.11	0.09	0.09	0.26	0.23	0.23	0.21	—	0.31	0.11	0.11	0.10	0.10
9. <i>Chlamydia (Genus)</i>	0.08	0.11	0.06	0.21	0.19	0.21	0.21	0.31	—	0.14	0.12	0.10	0.12
10. <i>Methanobacterium thermoautotrophicum</i>	0.11	0.10	0.10	0.11	0.06	0.11	0.12	0.11	0.14	—	0.51	0.25	0.20
11. <i>M. thermoautotrophicum strain M3</i>	0.11	0.10	0.10	0.12	0.07	0.12	0.12	0.12	0.12	0.51	—	0.25	0.24
12. <i>Methanobacterium sp. Carotococcus JR-1</i>	0.08	0.12	0.09	0.07	0.06	0.06	0.09	0.10	0.10	0.25	0.25	—	0.32
13. <i>Methanopyrus kandleri</i>	0.08	0.07	0.07	0.12	0.08	0.12	0.10	0.12	0.20	0.24	0.32	0.32	—

RESULTS

08F08C - 24 Jul 2012 - 0 paper, #2012, R.J. Robble

RCN·4·GSC

Eu

The eukaryotes were all similar to each other, Range for all eu-eu comparisons: 0.29 - 0.36

08F08C - 24 Jul 2012 - 0 paper, #2012, R.J. Robble

RCN·4·GSC

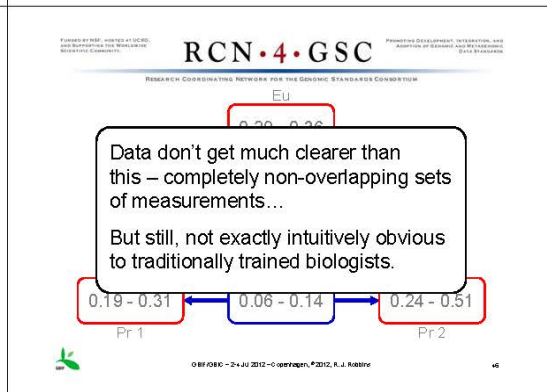
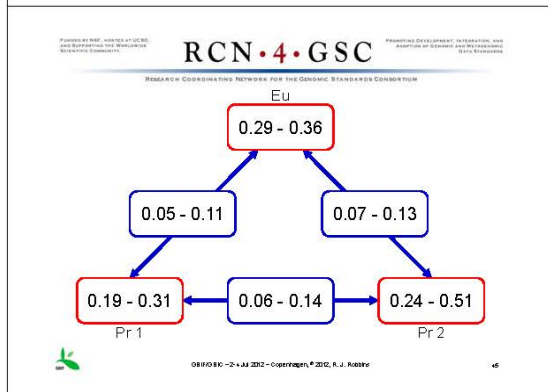
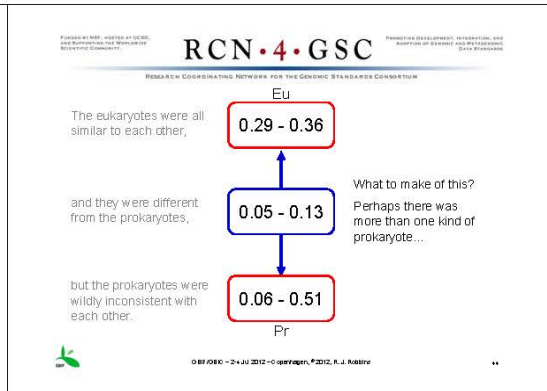
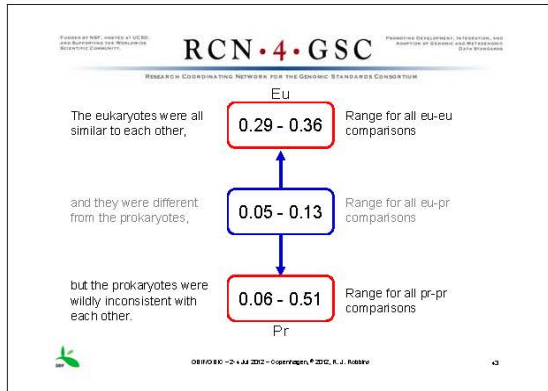
Eu

The eukaryotes were all similar to each other, Range for all eu-eu comparisons: 0.29 - 0.36

and they were different from the prokaryotes, Range for all eu-pr comparisons: 0.05 - 0.13

Pr

08F08C - 24 Jul 2012 - 0 paper, #2012, R.J. Robble



RCN·4·GSC

Aside:


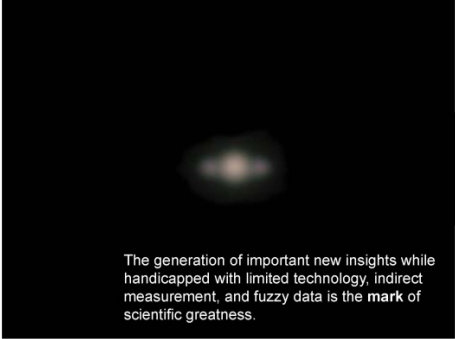




Most textbooks will tell you that, in 1610, Galileo Galilei became the first person to observe Saturn's rings.

But what did he really see?

OBFGSC - 24 Jul 2012 - Copenhagen, #2012, R.J. Robley





 <p>Or...</p>	 <p>The generation of important new insights while handicapped with limited technology, indirect measurement, and fuzzy data is the mark of scientific greatness.</p>
<p><small>Funded by NSF, HHS, and the European Union, and supported by the National Research Council</small></p> <p><b>RCN·4·GSC</b> <small>Promoting Development, Interaction, and Adoption of Standard and Interoperable Data Standards</small></p> <p><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p>What we can see is affected by (determined by?):</p> <p> GBFGBC – 2-4 Jul 2012 – Copenhagen, © 2012, R. J. Rizzo 51</p>	<p><small>Funded by NSF, HHS, and the European Union, and supported by the National Research Council</small></p> <p><b>RCN·4·GSC</b> <small>Promoting Development, Interaction, and Adoption of Standard and Interoperable Data Standards</small></p> <p><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p>What we can see is affected by (determined by?):</p> <p>Where we look.</p> <p> GBFGBC – 2-4 Jul 2012 – Copenhagen, © 2012, R. J. Rizzo 52</p>
<p><small>Funded by NSF, HHS, and the European Union, and supported by the National Research Council</small></p> <p><b>RCN·4·GSC</b> <small>Promoting Development, Interaction, and Adoption of Standard and Interoperable Data Standards</small></p> <p><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p>What we can see is affected by (determined by?):</p> <p>Where we look.</p> <p>The “illumination” available to us. actual lighting, instruments, analytical methods, other tools, ...</p> <p> GBFGBC – 2-4 Jul 2012 – Copenhagen, © 2012, R. J. Rizzo 53</p>	<p><small>Funded by NSF, HHS, and the European Union, and supported by the National Research Council</small></p> <p><b>RCN·4·GSC</b> <small>Promoting Development, Interaction, and Adoption of Standard and Interoperable Data Standards</small></p> <p><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p>What we can see is affected by (determined by?):</p> <p>Where we look.</p> <p>The “illumination” available to us. actual lighting, instruments, analytical methods, other tools, ...</p> <p>What we expect to see our theories, past experiences, biases, prejudices, ...</p> <p> GBFGBC – 2-4 Jul 2012 – Copenhagen, © 2012, R. J. Rizzo 54</p>

What we make of what we see is affected by:

Our ability to appreciate the details.

Our ability to see the big picture.

The context (vision) of our approach .

Our creativity.

.....



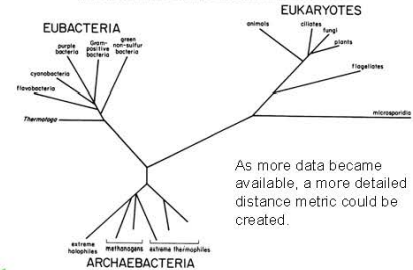
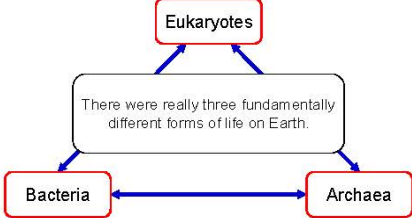
Evolution: Woese and Fox

Proc. Natl. Acad. Sci. USA 74 (1977) 5089

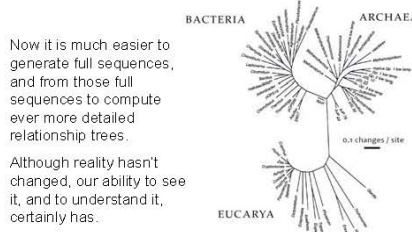
Table 1. Association coefficients (C<sub>ij</sub>) between representative members of the three primary kingdoms

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. <i>Staphylococcus aureus</i> , 18S	—	0.29	0.23	0.05	0.06	0.06	0.06	0.13	0.09	0.11	0.11	0.08	0.06
2. <i>Leucon micro</i> , 18S	0.29	—	0.28	0.10	0.06	0.06	0.10	0.09	0.11	0.10	0.13	0.07	0.07
3. <i>Leucon</i> , 18S	0.23	0.28	—	0.06	0.06	0.07	0.07	0.09	0.06	0.10	0.10	0.09	0.07
4. <i>Escherichia coli</i>	0.05	0.10	0.06	—	0.24	0.25	0.28	0.28	0.21	0.11	0.12	0.07	0.12
5. <i>Chlorella vulgaris</i> form	0.06	0.06	0.06	0.24	—	0.22	0.22	0.19	0.06	0.07	0.06	0.09	0.09
6. <i>Rhodospirillum rubrum</i>	0.06	0.06	0.07	0.22	0.22	—	0.24	0.26	0.20	0.11	0.11	0.08	0.12
7. <i>Corynebacterium diptheriae</i>	0.09	0.10	0.07	0.28	0.22	0.24	—	0.23	0.21	0.12	0.12	0.09	0.10
8. <i>Alphaproteobacteria</i> 17S	0.11	0.09	0.09	0.26	0.20	0.26	0.23	—	0.21	0.11	0.11	0.10	0.10

With his early tools, the best Woese could see were tables of laboriously created association coefficients..., but the implications were huge.

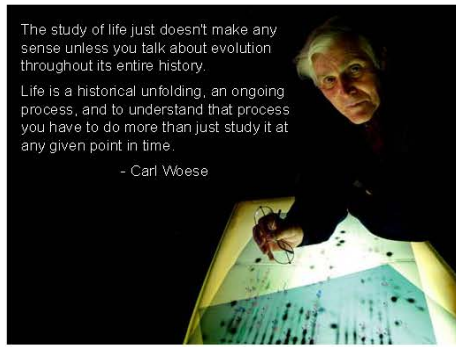


As more data became available, a more detailed distance metric could be created.



Now it is much easier to generate full sequences, and from those full sequences to compute ever more detailed relationship trees.

Although reality hasn't changed, our ability to see it, and to understand it, certainly has.



The study of life just doesn't make any sense unless you talk about evolution throughout its entire history.

Life is a historical unfolding, an ongoing process, and to understand that process you have to do more than just study it at any given point in time.

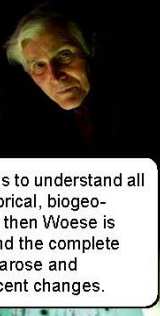
- Carl Woese



The study of life just doesn't make any sense unless you talk about evolution throughout its entire history.

Life is a historical unfolding, an ongoing process, and to understand that process you have to do more than just study it at any given point in time.


If the goal of biodiversity studies is to understand all life in the biosphere in its full historical, biogeographical, and functional context, then Woese is right. We must strive to understand the complete evolutionary context in which life arose and diversified, not merely catalog recent changes.



RCN • 4 • GSC

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Darwin knew that his model required some hereditary mechanism that could supply the variation upon which selection could work, but which would also be resistant to dilution through "blending." He never developed a working model of his own, and some of his provisional ideas flirted with Lamarckism.


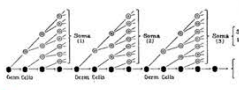


© 2012 RCN - 24 Jul 2012 - © Springer, #2012, R.J. Robbre

RCN • 4 • GSC

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

August Weismann's work on the germ-plasm theory assumed the hereditary stuff was in the cell nucleus and showed how this ruled out Lamarckian-style inheritance.





© 2012 RCN - 24 Jul 2012 - © Springer, #2012, R.J. Robbre

RCN • 4 • GSC

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

George Romanes thought Weismann's germ-plasm work ruling out the inheritance of acquired characteristics was an important extension to Darwin's own thinking, and so coined the phrase neo-Darwinism to describe this improved evolutionary model.




© 2012 RCN - 24 Jul 2012 - © Springer, #2012, R.J. Robbre

RCN • 4 • GSC

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Mendel's work could have provided the hereditary model, but his work was unknown to Darwin (and unknown to most of science) until 1900, when the rediscovery of his work triggered an explosion of new research, establishing the field of classical genetics.

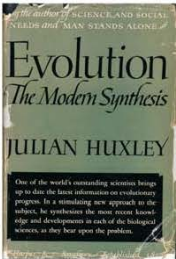


© 2012 RCN - 24 Jul 2012 - © Springer, #2012, R.J. Robbre

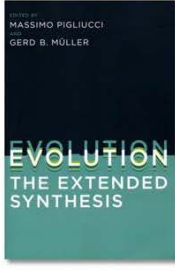
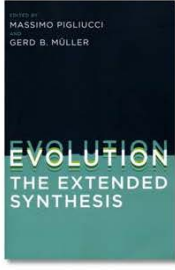
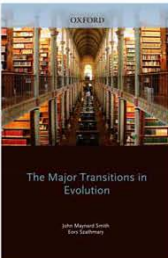
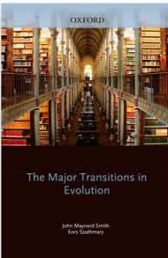
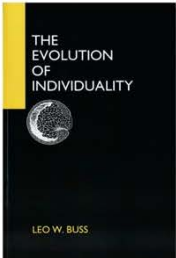
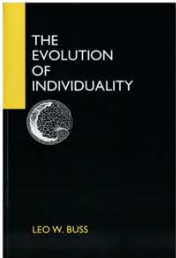
RCN • 4 • GSC

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM







The combination of neo-Darwinism and Mendelism produced *The Modern Synthesis* which has provided the intellectual foundation of most evolutionary thought from 1940 to the present.



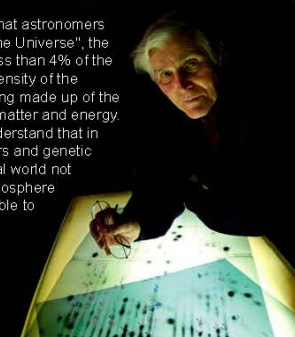
© 2012 RCN - 24 Jul 2012 - © Springer, #2012, R.J. Robbre

<p>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY RESEARCHER MEMBERSHIP COMMITTEES</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOLOGICAL DATA SCIENTISTS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <div style="border: 1px solid black; padding: 5px; margin: 10px 0;"> <p>Note that The Modern Synthesis was completed before Watson and Crick worked out the structure of DNA and before any tools of molecular biology were available to address problems of heredity, development, or evolution.</p> </div>  <p>© 2012 GSC - 24 JUL 2012 - © Springer, #2012, R.J. Raboin</p>	<p>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY RESEARCHER MEMBERSHIP COMMITTEES</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOLOGICAL DATA SCIENTISTS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <p>Today, some researchers are attempting to integrate newer findings from genomics and other fields to yield an improved and extended synthesis, suitable for 21<sup>st</sup>-century biology.</p>  <p>© 2012 GSC - 24 JUL 2012 - © Springer, #2012, R.J. Raboin</p>
<p>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY RESEARCHER MEMBERSHIP COMMITTEES</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOLOGICAL DATA SCIENTISTS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <div style="border: 1px solid black; padding: 5px; margin: 10px 0;"> <p>In planning for the future, GBIF would be well advised to attend carefully to these newly emerging evolutionary concepts.</p> <p>Evidence for new complexities and subtleties is growing, while some earlier fundamental assumptions are proving to be wrong.</p> <p>The possibility of significant extensions to our basic notions of organism and species seems not far off.</p> </div>  <p>© 2012 GSC - 24 JUL 2012 - © Springer, #2012, R.J. Raboin</p>	<p>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY RESEARCHER MEMBERSHIP COMMITTEES</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOLOGICAL DATA SCIENTISTS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p>  <p>In the history of life on Earth, several major transitions have occurred.</p> <p>These transitions were significant enough to change the nature of the evolutionary process itself, making it impossible to apply assumptions and analyses from one side of a transition to the other.</p> <p>© 2012 GSC - 24 JUL 2012 - © Springer, #2012, R.J. Raboin</p>
<p>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY RESEARCHER MEMBERSHIP COMMITTEES</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOLOGICAL DATA SCIENTISTS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <div style="border: 1px solid black; padding: 5px; margin: 10px 0;"> <p>Replicating molecules → Populations of molecules</p> <p>Independent replicators → Chromosomes</p> <p style="padding-left: 40px;">RNA → DNA</p> <p style="padding-left: 40px;">Prokaryotes → Eukaryotes</p> <p style="padding-left: 40px;">Asexual clones → Sexual populations</p> <p style="padding-left: 40px;">Unicellularity → Multicellularity</p> <p style="padding-left: 40px;">Solitary individuals → Colonies</p> <p style="padding-left: 40px;">Primate societies → Human societies (language)</p> </div>  <p>© 2012 GSC - 24 JUL 2012 - © Springer, #2012, R.J. Raboin</p>	<p>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY RESEARCHER MEMBERSHIP COMMITTEES</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOLOGICAL DATA SCIENTISTS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <p>For more than 80% of the time life has been evolving on Earth, multicellular "individuals" did not exist.</p> <p>Even now, they occur in only a handful of top-level taxa.</p> <p>Thus, making the "individual" the centerpiece for understanding evolution and for classifying life on Earth seems problematic.</p>  <p>© 2012 GSC - 24 JUL 2012 - © Springer, #2012, R.J. Raboin</p>

<p> <small>FINANCED BY NSF, MERCK AT UCSD, AND SUPPORTED BY THE NATIONAL MICROBIAL COMMISSION</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p>For more than 80% of the</p> <div style="border: 1px solid black; padding: 5px;"> <p>Attempting to understand microbial communities by thinking of them as a bunch of little bitty mice is an activity that falls on a continuum somewhere between fruitless and just plain wrong.</p> </div> <p>standing and classifying life on Earth seems problematic.</p> <p style="text-align: right;"><b>LEO W. BUSS</b></p> <p><small>08F08C - 24 JUL 2012 - 0-sprafkin, #2012, R.J. Robins</small> <span style="float: right;">73</span></p>	<p> <small>FINANCED BY NSF, MERCK AT UCSD, AND SUPPORTED BY THE NATIONAL MICROBIAL COMMISSION</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"><b>Science = Light's Better</b></p> <p><small>08F08C - 24 JUL 2012 - 0-sprafkin, #2012, R.J. Robins</small> <span style="float: right;">74</span></p>
<p> <small>FINANCED BY NSF, MERCK AT UCSD, AND SUPPORTED BY THE NATIONAL MICROBIAL COMMISSION</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>Old Joke:</b></p> <p>A drunk is crawling around a lamp post on his hands and knees.</p> <p>A cop comes along ...</p> <p><b>Cop:</b> What are you doing?</p> <p><b>Drunk:</b> Looking for my car keys.</p> <p><b>Cop:</b> Are you sure you dropped them here?</p> <p><b>Drunk:</b> No, I dropped them in the alley.</p> <p><b>Cop:</b> So why are you looking here?</p> <p><b>Drunk:</b> Because the light's better.</p> <p><small>08F08C - 24 JUL 2012 - 0-sprafkin, #2012, R.J. Robins</small> <span style="float: right;">75</span></p>	<p> <small>FINANCED BY NSF, MERCK AT UCSD, AND SUPPORTED BY THE NATIONAL MICROBIAL COMMISSION</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>Old Joke:</b></p> <div style="border: 1px solid black; padding: 5px;"> <p>Science is a light's better endeavor in that research effort is not directed at areas where the work is technically infeasible. Research is directed where real, interpretable results may be obtained.</p> <p>We do, in fact, conduct research where the light's better.</p> <p>But, when the light changes, so does science.</p> <p>With better illumination, we look in new areas.</p> <p>We find new things...</p> </div> <p><b>Drunk:</b> Because the light's better.</p> <p><small>08F08C - 24 JUL 2012 - 0-sprafkin, #2012, R.J. Robins</small> <span style="float: right;">76</span></p>
<p> <small>FINANCED BY NSF, MERCK AT UCSD, AND SUPPORTED BY THE NATIONAL MICROBIAL COMMISSION</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>Old Joke:</b></p> <div style="border: 1px solid black; padding: 5px;"> <p>Science is a light's better endeavor in that research effort is not directed at areas where the work is technically infeasible. Research is directed where real, interpretable results may be obtained.</p> <p>We do, in fact, conduct research where the light's better.</p> <p>But, when the light changes, so does science.</p> <p>With better illumination, we look in new areas.</p> <p>We find new things...</p> </div> <p><b>Drunk:</b> Because <b>The light IS changing...</b></p> <p><small>08F08C - 24 JUL 2012 - 0-sprafkin, #2012, R.J. Robins</small> <span style="float: right;">77</span></p>	<p> <small>FINANCED BY NSF, MERCK AT UCSD, AND SUPPORTED BY THE NATIONAL MICROBIAL COMMISSION</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p><b>Before the Light Changed:</b></p> <p style="text-align: center;"><b>Biochemistry</b></p> <p>If the genes are conceived as chemical substances, only one class of compounds need be given to which they can be reckoned as belonging, and that is the proteins in the wider sense, on account of the inexhaustible possibilities for variation which they offer. ... Such being the case, the most likely role for the nucleic acids seems to be that of the structure-determining supporting substance.</p> <div style="border: 1px solid black; padding: 2px; margin-top: 10px;"> <p><small>T. Caspersson: 1936. Über den chemischen Aufbau der S-Strukturen des Zellkerns. <i>Acta Med. Scand.</i>, 73, Suppl. 8, 1-151.</small></p> </div> <p><small>08F08C - 24 JUL 2012 - 0-sprafkin, #2012, R.J. Robins</small> <span style="float: right;">78</span></p>

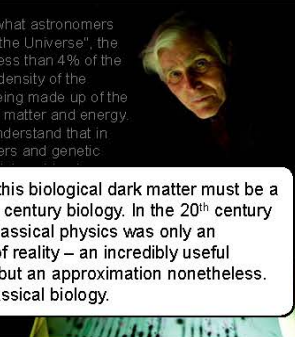
<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p>Before the Light Changed:</p> <p><b>Classical Genetics</b></p> <p>The genes are arranged [on chromosomes] in a manner similar to beads strung on a loose string.</p> <div style="border: 1px solid black; padding: 2px; width: fit-content; margin: 10px auto;"> <small>Sturtevant, A.H., and Beadle, G.W., 1939, <i>An Introduction to Genetics</i>. W.B. Saunders Company, Philadelphia, p. 94.</small> </div> <p>  <small>08F080C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robins</small> <span style="float: right;"><small>70</small></span> </p>	<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p>Before the Light Changed:</p> <p><b>Biodiversity</b></p> <p style="text-align: center; font-size: 2em;">???</p> <p>  <small>08F080C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robins</small> <span style="float: right;"><small>80</small></span> </p>
<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p>Before the Light Changed:</p> <p><b>Biodiversity</b></p> <p>Candidates:</p> <p>The centrality of the individual organism – the specimen – to biodiversity thinking.</p> <p>The centrality of the single-rooted tree of life as a device for representing our understanding of how life arose and diversified.</p> <p>  <small>08F080C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robins</small> <span style="float: right;"><small>81</small></span> </p>	<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"><b>Biological Dark Matter</b></p> <p style="text-align: center;"><b>Heaves into View</b></p> <p>  <small>08F080C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robins</small> <span style="float: right;"><small>82</small></span> </p>
<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"><b>Biological Dark Matter</b></p> <p style="text-align: center;"><b>Heaves into View</b></p> <p style="text-align: center; color: red;"><b>and what a view !</b></p> <p>  <small>08F080C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robins</small> <span style="float: right;"><small>83</small></span> </p>	<p> <small>FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL</small>  <b>RCN·4·GSC</b>  <small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small>  <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <ul style="list-style-type: none"> <li>• Metagenomics tools are showing that, compared to macro-scale organisms, the diversity of microbial communities is staggering. <ul style="list-style-type: none"> <li>• Intra-species bacterial genetic diversity is greater than that among the great apes; intra-genus bacterial diversity is greater than that among all the mammals.</li> </ul> </li> <li>• Metagenomics tools are showing that a full understanding of macro-scale organisms will depend on an understanding of their interactions with their associated microbiomes. <ul style="list-style-type: none"> <li>• Understanding how different ants optimize nutrient acquisition, and thus how they function in their niches, depends on an understanding of their associated gut microbiomes.</li> </ul> </li> </ul> <p>  <small>08F080C - 24 JUL 2012 - 0 openragen, #2012, R.J. Robins</small> <span style="float: right;"><small>84</small></span> </p>

We now know that what astronomers used to think of as "the Universe", the visible universe, is less than 4% of the total matter/energy density of the universe, the rest being made up of the still mysterious dark matter and energy. Similarly, we now understand that in terms of both numbers and genetic diversity, the microbial world not only dominates the biosphere but is almost impossible to sample properly.



We now know that what astronomers used to think of as "the Universe", the visible universe, is less than 4% of the total matter/energy density of the universe, the rest being made up of the still mysterious dark matter and energy. Similarly, we now understand that in terms of both numbers and genetic diversity, the microbial world not only dominates the biosphere but is almost impossible to sample properly.

Understanding this biological dark matter must be a top goal for 21<sup>st</sup> century biology. In the 20<sup>th</sup> century we found that classical physics was only an approximation of reality – an incredibly useful approximation, but an approximation nonetheless. So, too, with classical biology.



RCN • 4 • GSC  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Reality is NOT Negotiable, I  
 Genomics Example

© 2012 RCN • 4 • GSC

RCN • 4 • GSC  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM


Evolving Definition of a Gene:

- Classical Definition:** fundamental unit of heredity, mutation, and recombination (beads on a string).
- Physiological Definition:** fundamental unit of function (one gene, one enzyme).
- Cistronic Definition:** fundamental unit of expression (cis-trans test).
- Sequence Definition:** the smallest segment of the gene-string consistently associated with the occurrence of a specific genetic effect.
- Current Definition:** ???

© 2012 RCN • 4 • GSC

RCN • 4 • GSC  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Gene as Sequence (simplistic view)

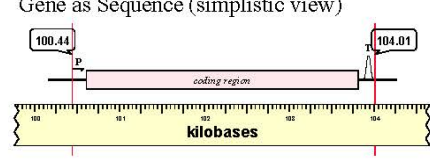


A gene is a transcribed region of DNA, flanked by upstream start regulatory sequences and downstream stop regulatory sequences.

© 2012 RCN • 4 • GSC

RCN • 4 • GSC  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Gene as Sequence (simplistic view)



The location of a gene can be designated by specifying the base-pair location of its beginning and end.

© 2012 RCN • 4 • GSC



FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENOMICS AND GENETICS
**RCN·4·GSC**
PROMOTING DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Simplistic View of Genome

DNA may be transcribed in either direction. Therefore, fully specifying a gene's position requires noting its orientation as well as its start and stop positions.

© 2012 GSC - 24 JUL 2012 - Openragen, #2012, R.J. Robine
51

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENOMICS AND GENETICS
**RCN·4·GSC**
PROMOTING DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Simplistic View of Genome

A naive view holds that a genome can be represented as a continuous linear string of nucleotides, with landmarks identified by the chromosome number followed by the offset number of the nucleotide at the beginning and end of the region of interest. This simplistic approach ignores the fact that human chromosomes may vary in length by tens of millions of nucleotides.

© 2012 GSC - 24 JUL 2012 - Openragen, #2012, R.J. Robine
52

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENOMICS AND GENETICS
**RCN·4·GSC**
PROMOTING DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Complicated Sequences

© 2012 GSC - 24 JUL 2012 - Openragen, #2012, R.J. Robine
53

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENOMICS AND GENETICS
**RCN·4·GSC**
PROMOTING DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Escherichia coli: the MMS Operon

Lupski, J.R., Goolson, G.N., 1989, DNA→DNA, and DNA→RNA→Protein: Orchestration by a single complex operon, *EvoEssays*, 10:152-157.

© 2012 GSC - 24 JUL 2012 - Openragen, #2012, R.J. Robine
54

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENOMICS AND GENETICS
**RCN·4·GSC**
PROMOTING DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Introns: Gart

© 2012 GSC - 24 JUL 2012 - Openragen, #2012, R.J. Robine
55

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENOMICS AND GENETICS
**RCN·4·GSC**
PROMOTING DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Introns: Gart

© 2012 GSC - 24 JUL 2012 - Openragen, #2012, R.J. Robine
56

<p>FINANCED BY NSF, NIMH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT</p> <p style="text-align: center;"><b>RCN • 4 • GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <p><b>Introns: Gart</b></p> <p style="text-align: right;">57</p>	<p>FINANCED BY NSF, NIMH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT</p> <p style="text-align: center;"><b>RCN • 4 • GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <p><b>Introns: Gart</b></p> <p style="text-align: right;">58</p>
<p>FINANCED BY NSF, NIMH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT</p> <p style="text-align: center;"><b>RCN • 4 • GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <p><b>Alternative Splicing: Gart</b></p> <p style="text-align: right;">59</p>	<p>FINANCED BY NSF, NIMH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT</p> <p style="text-align: center;"><b>RCN • 4 • GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <p><b>Nested Genes: Gart/Lcp</b></p> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> <p>Henikoff, S., Korte, M.A., Fochl, K., and Feistrom, J.W., 1986. Genes within a gene: Nested <i>Drosophila</i> genes encode unrelated proteins on opposite strands, <i>Cell</i> 44:33.</p> </div> <p style="text-align: right;">60</p>
<p>FINANCED BY NSF, NIMH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT</p> <p style="text-align: center;"><b>RCN • 4 • GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <p><b>Nested Gene Families: UGT1</b></p> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> <p>Ritter, J.K., Chen, F., et al., 1992. A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini, <i>J. Biol. Chem.</i> 267:3257.</p> </div> <p style="text-align: right;">61</p>	<p>FINANCED BY NSF, NIMH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT</p> <p style="text-align: center;"><b>RCN • 4 • GSC</b></p> <p style="text-align: center;">PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</p> <p style="text-align: center;">RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</p> <p><b>Nested Gene Families: UGT1</b></p> <p>phenol UDP-glucuronosyltransferase:</p> <p>bilirubin UDP-glucuronosyltransferase:</p> <p style="text-align: right;">62</p>

FUNDED BY NSF, NIH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENES AND REGULATORY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Nested Gene Families: UGT1

phenol UDP-glucuronosyltransferase:

bilirubin UDP-glucuronosyltransferases:

I challenge anyone to try to produce a definition of a gene that makes sense here.

081030C - 24 JUL 2012 - Cooperagen, #2012, R.J. Robins
103

FUNDED BY NSF, NIH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENES AND REGULATORY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Multiple Gene Products: POMC

Preproopiomelanocortin

Depending upon the tissue in which the gene is expressed, the POMC locus yields different protein products through alternative processing of the resulting polypeptide.

081030C - 24 JUL 2012 - Cooperagen, #2012, R.J. Robins
104

FUNDED BY NSF, NIH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENES AND REGULATORY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### Multiple Gene Products: POMC

Preproopiomelanocortin

One gene, one polypeptide?

Not quite. More like one gene, a dozen polypeptides, more or less, depending...

Depending upon the tissue in which the gene is expressed, the POMC locus yields different protein products through alternative processing of the resulting polypeptide.

081030C - 24 JUL 2012 - Cooperagen, #2012, R.J. Robins
105

FUNDED BY NSF, NIH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENES AND REGULATORY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

*Before Molecular Biology:*

Genes are the fundamental units of mutation, recombination, and heredity; they are arranged on the chromosomes like beads on a string.

081030C - 24 JUL 2012 - Cooperagen, #2012, R.J. Robins
106

FUNDED BY NSF, NIH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENES AND REGULATORY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

*Before Molecular Biology:*

Genes are the fundamental units of mutation, recombination, and heredity; they are arranged on the chromosomes like beads on a string.

*After Molecular Biology:*

No fundamental units, no beads, and no string...

081030C - 24 JUL 2012 - Cooperagen, #2012, R.J. Robins
107

FUNDED BY NSF, NIH, AND UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON GENETICS AND DEVELOPMENT
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENES AND REGULATORY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

## Reality is NOT Negotiable, II

## Biodiversity

081030C - 24 JUL 2012 - Cooperagen, #2012, R.J. Robins
108



FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

## The Tree of Life

© 2012 GSC - 24 JUL 2012 - OpenStax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

**Darwin's Tree**

© 2012 GSC - 24 JUL 2012 - OpenStax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

**Haeckel's Paleontology Tree**

© 2012 GSC - 24 JUL 2012 - OpenStax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

**Haeckel's Monophylogeny Tree**

© 2012 GSC - 24 JUL 2012 - OpenStax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

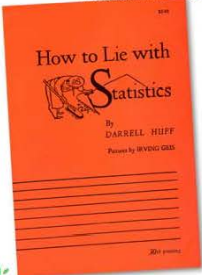


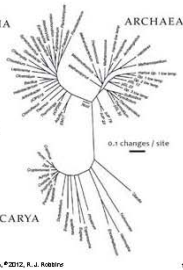
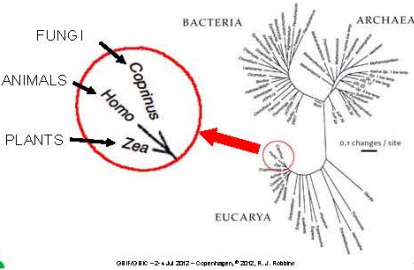
**A Pretty Tree**

© 2012 GSC - 24 JUL 2012 - OpenStax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF GENOMIC AND BIOINFORMATIC DATA STANDARDS  
RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

## Can Trees Lie?

© 2012 GSC - 24 JUL 2012 - OpenStax, #2012, R.J. Robble

<p>RCN·4·GSC</p>  <p>Anyone of the right age probably remembers this book and how effectively Huff demonstrated the ease with which graphical devices can misrepresent quantitative information.</p> <p>For example, ...</p>	<p>RCN·4·GSC</p>  <p>Lots of diversity up here</p>
<p>RCN·4·GSC</p>  <p>Lots of diversity up here</p> <p>Not so much down here</p>	<p>RCN·4·GSC</p>  <p>Another tree, this one based on comparisons of rRNA short sub-unit sequences.</p> <p>Branch length reflects actual divergence of sequence.</p>
<p>RCN·4·GSC</p>  <p>FUNGI ANIMALS PLANTS</p> <p>BACTERIA ARCHAEA</p> <p>EUCARYA</p>	<p>RCN·4·GSC</p> <div style="border: 1px solid black; padding: 10px;"> <p>If a tree is drawn to reflect <b>physiological</b> diversity, all of the differences among plants, animals, and fungi barely qualify as variations on a theme.</p> <p>Mammals are essentially the same physiological trick, served up in a variety of different packages.</p> </div>

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOINFORMATIC DATA PROCESSES

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

**Average Mammalian Body Temperatures:**

Human	37° C
Baboon	38° C
Fur seal	38° C
Humpback whale	36° C
Mouse ( <i>Mus musculus</i> )	37° C
Elephant	36° C
Polar Bear	37° C

© 2012 RCN - 24 Jul 2012 - © openstax, #2012, R.J. Robitre

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOINFORMATIC DATA PROCESSES

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM


**When is a Tree not a Tree?**

© 2012 RCN - 24 Jul 2012 - © openstax, #2012, R.J. Robitre

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOINFORMATIC DATA PROCESSES

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

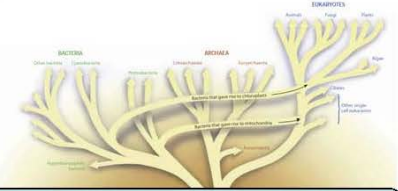
**Most multicellular eukaryotic taxa can be arranged in a tree-like configuration, but when we include the origin of intra-cellular organelles things get more complicated...**



© 2012 RCN - 24 Jul 2012 - © openstax, #2012, R.J. Robitre

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOINFORMATIC DATA PROCESSES

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM



**The origin of mitochondria, chloroplasts, and several other eukaryotic cell inclusions through endosymbiosis means that, technically speaking at least, not only are eukaryotic taxa polyphyletic, so are eukaryotic "individuals".**

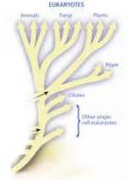
© 2012 RCN - 24 Jul 2012 - © openstax, #2012, R.J. Robitre

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOINFORMATIC DATA PROCESSES

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

**In addition...**

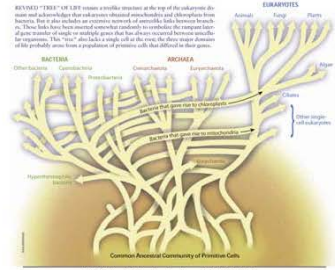
**We now know that horizontal gene transfer (HGT) occurs regularly among prokaryotes and most likely was the dominant form of inheritance during the early evolution of life on Earth.**



© 2012 RCN - 24 Jul 2012 - © openstax, #2012, R.J. Robitre

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING RESEARCHING COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND BIOINFORMATIC DATA PROCESSES

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM



© 2012 RCN - 24 Jul 2012 - © openstax, #2012, R.J. Robitre

RCN·4·GSC

When trying to represent the evolutionary history of ALL life on Earth, a single-rooted tree is a poor data model.

© 2012 R.J. Robley

RCN·4·GSC

Classical biology has also saddled us with a phylogenetic tree, an image the biologist invests with a deep and totally unwarranted significance. The tree is no more than a representational device, but to the biologist it is some God-given truth. Thus, for example, we agonize over how the tree can accommodate horizontal gene transfer events, when it should simply be a matter of when (and to what extent) the evolution course can be usefully represented by a tree diagram.

Evolution defines the tree, not the reverse.

© 2012 R.J. Robley

RCN·4·GSC

## The Individual

© 2012 R.J. Robley

RCN·4·GSC

The traditional biologist's view of life, first stated by Aristotle, starts by looking at individual organisms and asking what properties they have in common. (p 5)

Mosconi, Harold J. 1992. *Beginnings of Cellular Life: Abiogenesis Speculates Biogenesis*. New Haven, Yale University Press.

© 2012 R.J. Robley

RCN·4·GSC

The subjects of classification are organisms and the subjects of taxonomy are classifications. (p11)

It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)

George Gaylord Simpson. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.

© 2012 R.J. Robley

RCN·4·GSC

### Weismann's Germ Plasm Theory

Organisms that best satisfy the notion of exhibiting "completely objective" individuals are animals that follow a Weismannian pattern of development – that is, an early sequestering of a separate germ line, with a complete logical and physical separation of somatic and germ tissue. Such animals begin as a zygote, then develop mitotically into a multicellular adult that, with luck, lives to adulthood and reproduces via the meiotic production of gametes.

© 2012 R.J. Robley

<p>RCN·4·GSC</p> <p>environment</p> <p>Inputs: food, water, gases → organism (genome, F<sub>1</sub>, F<sub>2</sub>, F<sub>n</sub>)</p> <p>Outputs: wastes, artifacts, gases ← organism</p> <p>Such an individual can, with a certain amount of conceptual legerdemain, be envisioned as functioning autonomously in an environment, from which it acquires a few necessary inputs and into which it delivers certain outputs. It has a distinct body, with a clear external-internal boundary. Its phenotype is determined by its genotype, with internal physiological functions being carried out under the protein-mediated instructions of the genome. Reproduction involves "generations" in which new individuals come into existence, mature, and die.</p> <p>© 2012 R.J. Robley</p>	<p>RCN·4·GSC</p> <p>environment</p> <p>Inputs: food, water, gases → organism (genome, F<sub>1</sub>, F<sub>2</sub>, F<sub>n</sub>)</p> <p>Outputs: wastes, artifacts, gases ← organism</p> <p>REPRODUCTION / DEVELOPMENT</p> <p>organism (genome, F<sub>1</sub>, F<sub>2</sub>, F<sub>n</sub>)</p> <p>© 2012 R.J. Robley</p>
<p>RCN·4·GSC</p> <p>The subjects of classification are organisms and the subjects of taxonomy are classifications. (p11)</p> <p>It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)</p> <p>George Gaylord Simpson. 1961. <i>Principles of Animal Taxonomy</i>. New York: Columbia University Press.</p> <p>© 2012 R.J. Robley</p>	<p>RCN·4·GSC</p> <p>The subjects of classification are organisms and the subjects of taxonomy are classifications. (p11)</p> <p>Not very many...</p> <p>It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)</p> <p>George Gaylord Simpson. 1961. <i>Principles of Animal Taxonomy</i>. New York: Columbia University Press.</p> <p>© 2012 R.J. Robley</p>
<p>RCN·4·GSC</p> <p>The subjects of classification are organisms and the subjects of taxonomy are classifications. (p11)</p> <p>Not very important...</p> <p>It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)</p> <p>George Gaylord Simpson. 1961. <i>Principles of Animal Taxonomy</i>. New York: Columbia University Press.</p> <p>© 2012 R.J. Robley</p>	<p>RCN·4·GSC</p> <p>The subjects of classification are organisms and the subjects of taxonomy are classifications. (p11)</p> <p>Without doubt... Indubitably... Really ? ! ?</p> <p>It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)</p> <p>George Gaylord Simpson. 1961. <i>Principles of Animal Taxonomy</i>. New York: Columbia University Press.</p> <p>© 2012 R.J. Robley</p>



FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

### What about lichens?

It is not easy to apply the concept of individual organism to a lichen.

08F080 - 24 JUL 2012 - Cooper, R.J. Robins
139

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Although lichens are composite structures, most are highly organized with distinctive morphologies.

Without a detailed microscopical examination, most lichens appear to be single entities.

08F080 - 24 JUL 2012 - Cooper, R.J. Robins
140

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

At the cellular level, it is possible to detect the presence of algae, embedded in the mycobiont tissue.

[L]ichens are not simple plants, not ordinary individuals in the ordinary sense of the word, they are, rather, colonies, which consist of hundreds or thousands of individuals, of which, however, one alone plays the master, while the rest, forever imprisoned, prepare the nutriment for themselves and their master. This fungus is a fungus of the class Ascomycetes, a parasite which is accustomed to live upon others' work. Its slaves are green algae, which it has sought out, or indeed caught hold of, and compelled into its service. It surrounds them as a spider its prey, with a fibrous net of narrow meshes, which is gradually converted into an impenetrable covering, but while the spider sucks its prey and leaves it dead, the fungus notes the algae found in its net to more rapid activity, even to more vigorous increase.

Schwendener (1899)

08F080 - 24 JUL 2012 - Cooper, R.J. Robins
141

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

At the cellular level, it is possible to detect the presence of algae, embedded in the mycobiont tissue.

Schwendener's prose is dramatic, but makes the key point: a lichen is a living unit, but is not an "individual" as classically conceived. Nor can it be decomposed into individuals without giving up both its essence and its viability.

A lichen is a composite organism that cannot be subdivided into "individuals" and remain living. How does this square with the idea of the "individual" as the true "fundamental unit" of nature?

08F080 - 24 JUL 2012 - Cooper, R.J. Robins
142

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

The same fungal partner can combine with different algal partners, to produce phenotypically distinct lichens, that occupy different niches.

In the right circumstances, they may be found in a gradient.

Are these two different lichens or ...

08F080 - 24 JUL 2012 - Cooper, R.J. Robins
143

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

The same fungal partner can combine with different algal partners, to produce phenotypically distinct lichens.

In the right circumstances, they may be found in a gradient.

Are these two different lichens or ...

Maybe lichens are just a marginal case – the exception that proves the rule, and all that ...


08F080 - 24 JUL 2012 - Cooper, R.J. Robins
144



FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

But they are certainly not rare...



© 2012 GSC - 24 JUL 2012 - © openstax, #2012, R.J. Robble


FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

What about termites?

They are a critically important, sometimes dominant, species in many ecosystems, yet they cannot exist without their gut symbionts.


Can we really dismiss termites as a rare, marginal case?



© 2012 GSC - 24 JUL 2012 - © openstax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM



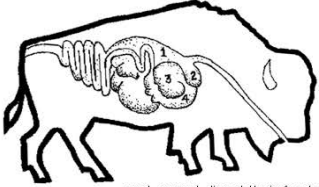
How about mammals? Surely they exemplify the idealized rugged individualism of autonomous organisms – the fundamental (and completely objective) unit of nature...

© 2012 GSC - 24 JUL 2012 - © openstax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

But wait, bison are ruminants



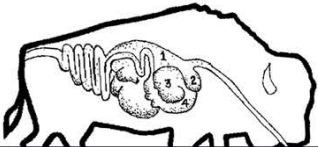
and cannot digest their food without the assistance of their gut flora.

© 2012 GSC - 24 JUL 2012 - © openstax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

But wait, bison are ruminants



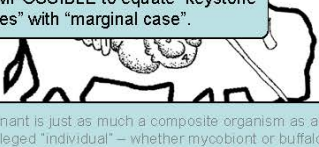
A ruminant is just as much a composite organism as a lichen. The alleged "individual" – whether mycobiont or buffalo – cannot obtain nutrients, and thus cannot live, without its microbial partners.

© 2012 GSC - 24 JUL 2012 - © openstax, #2012, R.J. Robble

FUNDED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY THE NATIONAL RESEARCH COUNCIL ON SCIENTIFIC COMMUNITY
**RCN·4·GSC**
PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF SCIENTIFIC AND EDUCATIONAL DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Ruminants are keystone species in many grassland ecosystems. It is IMPOSSIBLE to equate "keystone species" with "marginal case".



A ruminant is just as much a composite organism as a lichen. The alleged "individual" – whether mycobiont or buffalo – cannot obtain nutrients, and thus cannot live, without its microbial partners.

© 2012 GSC - 24 JUL 2012 - © openstax, #2012, R.J. Robble

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING COMMUNITY

**RCN·4·GSC**

PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF OPENNESS AND REPRODUCIBILITY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM



OK, people then.  
Surely human beings aren't just hopped up lichens...

0816080 - 24 JUL 2012 - 0openragen, #2012, R.J. Robbin

151

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING COMMUNITY

**RCN·4·GSC**

PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF OPENNESS AND REPRODUCIBILITY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM



8 June 2012

14 June 2012

0816080 - 24 JUL 2012 - 0openragen, #2012, R.J. Robbin

152

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING COMMUNITY

**RCN·4·GSC**

PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF OPENNESS AND REPRODUCIBILITY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM



0816080 - 24 JUL 2012 - 0openragen, #2012, R.J. Robbin

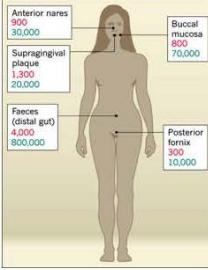
153

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING COMMUNITY

**RCN·4·GSC**

PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF OPENNESS AND REPRODUCIBILITY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM



Anterior nares	900	Buccal mucosa	800
Supragingival plaque	1,300		70,000
Faeces (distal gut)	4,000		
	800,000	Posterior fornx	300
			10,000

Fig 1 Variation in diversity. Researchers of the Human Microbiome Project are studying the microbial inhabitants of the human body using samples taken from 242 healthy adults at 15 (for males) or 18 (for female) body sites — from the skin (our armpits), mouth and throat (one also), vagina (three sites), nostrils and faeces to represent the distal gastrointestinal tract). Huttenhower et al. and Meeks et al. have estimated the number of microbial species and their genes in these samples, and found substantial variation in microbial community composition at different body habitats. The two groups used different counting methodologies, and their numbers vary accordingly, such that exact figures are not available. However, crude estimates of number of microbial species (red) and a number of microbial genes (blue) are shown for examples of sites containing high species diversity, such as the gastrointestinal tract and teeth (supragingival plaque), sites with intermediate diversity such as the nose of the cheek (buccal mucosa) and nostrils (anterior nares), and sites with lower diversity such as the vaginal posterior fornx. The authors also found substantial variation in both the diversity and the composition of the microbial communities at different sites within the same general body region.

0816080 - 24 JUL 2012 - 0openragen, #2012, R.J. Robbin

154

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING COMMUNITY

**RCN·4·GSC**

PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF OPENNESS AND REPRODUCIBILITY DATA STANDARDS

RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM

Microbial World:

- In addition to being ubiquitous and abundant on and in every macroscale organism, prokaryotes occur in every imaginable environment (and maybe a few more).

0816080 - 24 JUL 2012 - 0openragen, #2012, R.J. Robbin

155

FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTING THE RESEARCHING COMMUNITY

**RCN·4·GSC**

PROMOTES DEVELOPMENT, INTEGRATION, AND ADOPTION OF OPENNESS AND REPRODUCIBILITY DATA STANDARDS







RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM



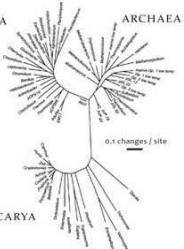

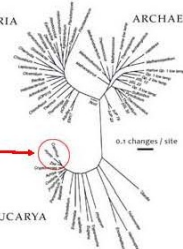

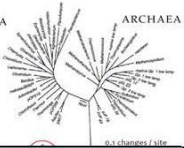






Microbial World:

- In addition to being ubiquitous and abundant on and in every macroscale organism, prokaryotes occur in every imaginable environment (and maybe a few more).
- They are abundant (more bacteria in a bucket of seawater than there are mammals in Africa).

0816080 - 24 JUL 2012 - 0openragen, #2012, R.J. Robbin

156

<p>Funded by NSF, NIH, and UCSD, and supported by the National Microbial Community.</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: right;"><small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small></p> <p><b>Microbial World:</b></p> <ul style="list-style-type: none"> <li>In addition to being ubiquitous and abundant on and in every macroscale organism, prokaryotes occur in every imaginable environment (and maybe a few more).</li> <li>They are abundant (more bacteria in a bucket of seawater than there are mammals in Africa).</li> <li>They are locally diverse (1 g of soil contains <math>10^7</math>-<math>10^9</math> prokaryotic cells, with 2,000–18,000 different genomes).</li> </ul>  <p style="text-align: center;"><small>081010C - 24 JUL 2012 - Copenhagen, #2012, R.J. Robins</small></p> <p style="text-align: right;"><small>157</small></p>	<p>Funded by NSF, NIH, and UCSD, and supported by the National Microbial Community.</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: right;"><small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small></p> <p><b>Microbial World:</b></p> <ul style="list-style-type: none"> <li>In addition to being ubiquitous and abundant</li> </ul> <div style="border: 1px solid black; padding: 5px; margin: 5px 0;"> <p>To repeat: That's 2,000 to 18,000 separate "species" in a teaspoon of soil...</p> </div> <p>of seawater than there are mammals in Africa).</p> <ul style="list-style-type: none"> <li>They are locally diverse (1 g of soil contains <math>10^7</math>-<math>10^9</math> prokaryotic cells, with 2,000–18,000 different genomes).</li> </ul>  <p style="text-align: center;"><small>081010C - 24 JUL 2012 - Copenhagen, #2012, R.J. Robins</small></p> <p style="text-align: right;"><small>158</small></p>
<p>Funded by NSF, NIH, and UCSD, and supported by the National Microbial Community.</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: right;"><small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small></p> <p><b>Bottom Line:</b></p> <ul style="list-style-type: none"> <li>This is fundamentally a microbial biosphere. Half of the biomass and most of the diversity occur in microbes.</li> </ul>  <p style="text-align: center;"><small>081010C - 24 JUL 2012 - Copenhagen, #2012, R.J. Robins</small></p> <p style="text-align: right;"><small>159</small></p>	<p>Funded by NSF, NIH, and UCSD, and supported by the National Microbial Community.</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: right;"><small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small></p> <p><b>Bottom Line:</b></p> <ul style="list-style-type: none"> <li>This is fundamentally a microbial biosphere. Half of the biomass and most of the diversity occur in microbes.</li> <li>Microbes occur in free-living communities and also in very tight, functional associated with ALL multi-cellular organisms.</li> </ul>  <p style="text-align: center;"><small>081010C - 24 JUL 2012 - Copenhagen, #2012, R.J. Robins</small></p> <p style="text-align: right;"><small>160</small></p>
<p>Funded by NSF, NIH, and UCSD, and supported by the National Microbial Community.</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: right;"><small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small></p> <p><b>Bottom Line:</b></p> <ul style="list-style-type: none"> <li>This is fundamentally a microbial biosphere. Half of the biomass and most of the diversity occur in microbes.</li> <li>Microbes occur in free-living communities and also in very tight, functional associated with ALL multi-cellular organisms.</li> <li>The notion of individual organisms, as fundamental units in nature, is not objective "truth" – instead it is, at best, a useful approximation.</li> </ul>  <p style="text-align: center;"><small>081010C - 24 JUL 2012 - Copenhagen, #2012, R.J. Robins</small></p> <p style="text-align: right;"><small>161</small></p>	<p>Funded by NSF, NIH, and UCSD, and supported by the National Microbial Community.</p> <p style="text-align: center;"><b>RCN·4·GSC</b></p> <p style="text-align: center;"><small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small></p> <p style="text-align: right;"><small>PROMOTES DEVELOPMENT, INTEGRATION, AND ADAPTATION OF GENOMIC AND METAGENOMIC DATA STANDARDS</small></p> <p><b>Bottom Line:</b></p> <div style="border: 1px solid black; padding: 5px; margin: 5px 0;"> <p>From the perspective of community biology (which arguably is synonymous with "biology"), the "individual" is a reductionist abstraction.</p> <p>It is useful in the way "assume a spherical cow" is useful in biophysics – it simplifies the analysis, but at some cost to its correspondence with reality.</p> </div> <p>mental units in nature, is not objective "truth" – instead it is, at best, a useful approximation.</p>  <p style="text-align: center;"><small>081010C - 24 JUL 2012 - Copenhagen, #2012, R.J. Robins</small></p> <p style="text-align: right;"><small>162</small></p>

<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY MEMBERS OF RESEARCHING COMMUNITY</small> </p> <p style="text-align: center;"> <b>RCN·4·GSC</b> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <small>PROMOTING DEVELOPMENT, INTERACTION, AND ASSISTANCE IN GENOMIC AND BIOINFORMATIC DATA STANDARDS</small> </p> <p>         The idea that the whole may be understood by understanding all of its parts is the conceit of reductionism.       </p> <p style="text-align: center;">  <small>08/08/08 - 24 JUL 2012 - Copenhagen, © 2012, R. J. Robbins</small> </p>	<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY MEMBERS OF RESEARCHING COMMUNITY</small> </p> <p style="text-align: center;"> <b>RCN·4·GSC</b> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <small>PROMOTING DEVELOPMENT, INTERACTION, AND ASSISTANCE IN GENOMIC AND BIOINFORMATIC DATA STANDARDS</small> </p> <p>         The idea that the whole may be understood by understanding all of its parts is the conceit of reductionism.       </p> <p>         The idea that the whole may be understood by understanding a few (or only one) of its parts is simply wrong.       </p> <p style="text-align: center;">  <small>08/08/08 - 24 JUL 2012 - Copenhagen, © 2012, R. J. Robbins</small> </p>
<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY MEMBERS OF RESEARCHING COMMUNITY</small> </p> <p style="text-align: center;"> <b>RCN·4·GSC</b> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <small>PROMOTING DEVELOPMENT, INTERACTION, AND ASSISTANCE IN GENOMIC AND BIOINFORMATIC DATA STANDARDS</small> </p> <p>         If the goal of biodiversity studies is to understand all of the diversity in the Earth's biosphere...       </p> <div style="display: flex; align-items: center;">  </div> <p style="text-align: center;">  <small>08/08/08 - 24 JUL 2012 - Copenhagen, © 2012, R. J. Robbins</small> </p>	<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY MEMBERS OF RESEARCHING COMMUNITY</small> </p> <p style="text-align: center;"> <b>RCN·4·GSC</b> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <small>PROMOTING DEVELOPMENT, INTERACTION, AND ASSISTANCE IN GENOMIC AND BIOINFORMATIC DATA STANDARDS</small> </p> <p>         If the goal of biodiversity studies is to understand all of the diversity in the Earth's biosphere...       </p> <div style="display: flex; align-items: center;">  </div> <p>         Then the notion that we can accomplish that goal only by looking here is just plain wrong.       </p> <p style="text-align: center;">  <small>08/08/08 - 24 JUL 2012 - Copenhagen, © 2012, R. J. Robbins</small> </p>
<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY MEMBERS OF RESEARCHING COMMUNITY</small> </p> <p style="text-align: center;"> <b>RCN·4·GSC</b> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <small>PROMOTING DEVELOPMENT, INTERACTION, AND ASSISTANCE IN GENOMIC AND BIOINFORMATIC DATA STANDARDS</small> </p> <p>         If the goal of biodiversity studies is to understand all of the diversity in the Earth's biosphere...       </p> <div style="display: flex; align-items: center;">  </div> <p>         Then the notion that we can accomplish that goal only by looking here...       </p> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> <p>         To maintain relevancy, to deal with the non-negotiable aspects of nature, 21<sup>st</sup> century biodiversity studies MUST include a large and growing commitment to understanding microbial biodiversity.       </p> </div> <p style="text-align: center;">  <small>08/08/08 - 24 JUL 2012 - Copenhagen, © 2012, R. J. Robbins</small> </p>	<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY MEMBERS OF RESEARCHING COMMUNITY</small> </p> <p style="text-align: center;"> <b>RCN·4·GSC</b> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <small>PROMOTING DEVELOPMENT, INTERACTION, AND ASSISTANCE IN GENOMIC AND BIOINFORMATIC DATA STANDARDS</small> </p> <p> <i>Before Genomics:</i>          Individual organisms are the fundamental units of biodiversity; their evolutionary history can be explained by arranging them into groups, with the groups composed into a single-rooted tree.       </p> <p style="text-align: center;">  <small>08/08/08 - 24 JUL 2012 - Copenhagen, © 2012, R. J. Robbins</small> </p>
<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY MEMBERS OF RESEARCHING COMMUNITY</small> </p> <p style="text-align: center;"> <b>RCN·4·GSC</b> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <small>PROMOTING DEVELOPMENT, INTERACTION, AND ASSISTANCE IN GENOMIC AND BIOINFORMATIC DATA STANDARDS</small> </p> <p> <i>Before Genomics:</i>          Individual organisms are the fundamental units of biodiversity; their evolutionary history can be explained by arranging them into groups, with the groups composed into a single-rooted tree.       </p> <p> <i>After Genomics</i>          No "completely objective" individuals, no one true tree.       </p> <p style="text-align: center;">  <small>08/08/08 - 24 JUL 2012 - Copenhagen, © 2012, R. J. Robbins</small> </p>	<p> <small>FINANCED BY NSF, HOSTED AT UCSD, AND SUPPORTED BY MEMBERS OF RESEARCHING COMMUNITY</small> </p> <p style="text-align: center;"> <b>RCN·4·GSC</b> </p> <p style="text-align: center;"> <small>RESEARCH COORDINATING NETWORK FOR THE GENOMIC STANDARDS CONSORTIUM</small> </p> <p style="text-align: center;"> <small>PROMOTING DEVELOPMENT, INTERACTION, AND ASSISTANCE IN GENOMIC AND BIOINFORMATIC DATA STANDARDS</small> </p> <p style="text-align: center;">         Robert J. Robbins          RJR8222@gmail.com       </p> <p style="text-align: center; font-size: 2em; font-weight: bold;">         END       </p> <p style="text-align: center;">  <a href="http://www.rj-robbins.com/slides/RJR-GBIC-2012.pdf">http://www.rj-robbins.com/slides/RJR-GBIC-2012.pdf</a>  </p> <p style="text-align: center;">  <small>08/08/08 - 24 JUL 2012 - Copenhagen, © 2012, R. J. Robbins</small> </p>



## 9 會議活動照片



大會團體照



分組會議中各領域科學家齊聚討論場景



小組討論場景

