

# Bioinformática

Nuevas Capacidades  
para un Nuevo País

---

**Carlos Vásquez**

Director Nacional de Tecnología y Competitividad  
Microsoft Colombia



La Pregunta del Millón de Dólares

1<sup>st</sup>

Country in  
Biodiversity per  
square Km

1<sup>st</sup>

Birds  
1855 species

2<sup>nd</sup>

Plants  
41000 species



3<sup>rd</sup>

Reptiles  
524 species

4<sup>th</sup>

Amphibian  
750 species

5<sup>th</sup>

Mammals  
471 species



**Biodiversity** is not only a rich natural resource, but also, an **strategic asset** for enhance **competitiveness** and economic development with **environmental sustainability**.





*"The **computer industry** knows that scientists can come up with **strange ideas** and requirements that may well, in time, have broader commercial application elsewhere. This is one of the reasons why Microsoft is engaging the scientific community with its new '**Towards 2020 Science**' report on computers in science. That report inspired this week's focus on computing in Nature."*

Steering the future of computing. [Nature 440\(7083\)](#)

2006

Microsoft  
**Research**

COLLABORATION • INNOVATION • INTEROPERABILITY • DISCOVERY • PARTNERSHIP

External Research Division



# 2007: Microsoft Research Tour







2007: Microsoft Research Tour

Microsoft  
Research  
www.microsoft.com





## Biodiversidad + Talento Humano

*“Los científicos de Microsoft Research se impresionaron con la alta calidad de las investigaciones que realizan algunas de las instituciones visitadas en Colombia [...]. Creo que hay una oportunidad única para convertir a Colombia en un modelo de investigación y desarrollo para toda la región [...]. Para alcanzar estas metas, que elevarían el país a un nuevo nivel de competitividad, es crucial establecer objetivos con respecto a inversiones en ciencia, investigación y desarrollo en concordancia con el Producto Interno Bruto del país”.*

**Bill Gates**

***Apartes de carta entregada al presidente Álvaro Uribe, después de la primera visita de Microsoft Research a Colombia.***

1. Fortalecer **Comunidad Científica Actual**
2. **Infraestructura Local Como Catalizador**
3. Crear **Capacidades** en Talento Local
4. Oportunidades en **La Industria Software**



# Nuevas Capacidades para un Nuevo País

1



Comunidad Científica

Microsoft® Research

# FacultySummit 2011

**Cartagena, Colombia | May 18–20 | In partnership with COLCIENCIAS**

Microsoft Research hosted its seventh annual Latin American Faculty Summit in Cartagena, Colombia, May 18 to 20, 2011. This event was organized in partnership with [Colciencias](#). The theme of the 2011 summit was *Advancing Computing; Advancing Science*. [View videos](#) of the presentations from the summit.



Cartagena, Colombia



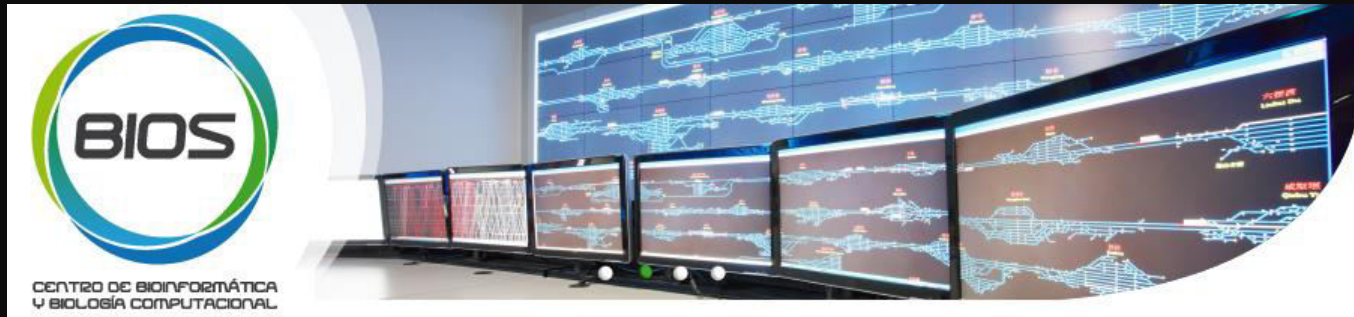
# Nuevas Capacidades para un Nuevo País



2

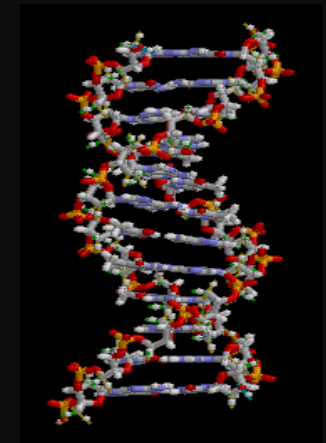
Infraestructura

# Centro Nacional de Bioinformática



#1

Biodiverse  
Country in the  
World per Km<sup>2</sup>





# 2013: BIOS – Centro Nacional de Bioinformática



# Nuevas Capacidades para un Nuevo País



3

Construir Capacidades





# Microsoft Research



[Our research](#) [Connections](#) [Careers](#) [About us](#)

[All](#) [Downloads](#) [Events](#) [Groups](#) [News](#) [People](#) [Projects](#) [Publications](#) [Videos](#)

## Biological Computation



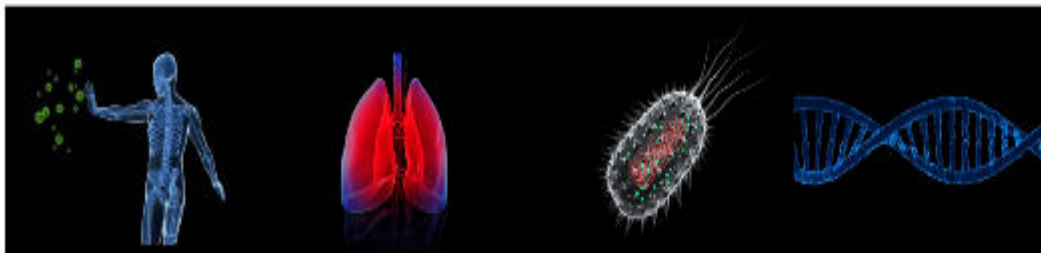
Our group is conducting research to uncover fundamental principles of biological computation: what biological systems compute, and how this computation is performed. We are also developing methods for programming computation in biological systems. Our research currently focuses on fundamentals of Biological Computation, with applications in Immunology and Development, together with principles of Programming Life, with applications in DNA Computing and Synthetic Biology.

Organism

Organ

Cell

Molecule



Immunology

Developmental Biology

Synthetic Biology

DNA Computing

[Biological Computation Home](#)

[Publications](#)

**We are NOW HIRING**



[Computational Science Lab](#)

- [Computational Science](#)
- [Biological Computation](#)
- [Computational Ecology and Environmental Science](#)

[Related Links](#)

- [The Microsoft Biology Initiative](#)

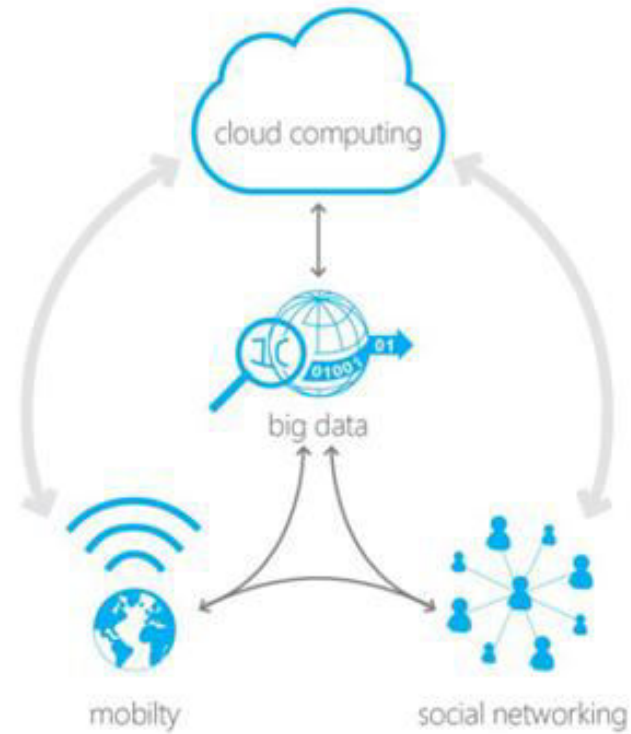


# Nuevas Capacidades para un Nuevo País



4

Industria Local de Software





# Microsoft Biology Initiative

Microsoft Biology Initiative

[.NET Bio \(formerly Microsoft Biology Foundation\)](#)

[Microsoft Biology Tool](#)



The Microsoft Biology Initiative (MBI) is a Microsoft Research effort to bring new technology and tools to bioinformatics and biology research. This initiative is comprised of two primary components, .NET Bio (formerly Microsoft Biology Foundation [MBF]) and the Microsoft Biology Tools (MBT).

## **.NET Bio (formerly Microsoft Biology Foundation)**

.NET Bio is a language-neutral bioinformatics toolkit built as an extension to the Microsoft .NET Framework—initially aimed at the area of Genomics research. Currently, it implements a range of parsers for common bioinformatics file formats; a range of algorithms for manipulating DNA, RNA, and protein sequences; and a set of connectors to biological web services such as NCBI BLAST. .NET Bio is available under an open source license, and executables, source code, demo applications, and documentation are freely downloadable from the Codeplex site. [Learn more...](#)

## **Microsoft Biology Tools**

The Microsoft Biology Tools (MBT) are a collection of tools that enable biology and bioinformatics researchers to be more productive in making scientific discoveries. The tools in this catalog are free and, when available, include links to the source code and open source collaboration projects. [Browse the catalog...](#)

File Home Insert Page Layout Formulas Data Review View Load Test .NET Bio Team

Import From Sequence Data Export To Sequence Alignment Select Aligner Sequence Assembly Assemble Select BLAST service Web service Charts Charting Operate on Genomic Intervals Genomic Intervals Venn Diagram NodeXL Configure Configuration Home Page Help About .NET Bio

N3 fx T

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	
1	Sequence Data	C	C	T	T	T	C	C	C	A	G	G	G	A	C	T	T	C	T	A	C	A	A	G	G	A	A	A	A	A	G	C	T	A	G	A	G	T	T	G	G	T	T	A	C	T	G	A	C	T	T	C	
2		G	T	T	A	A	A	A	G	T	T	G	A	C	A	T	A	A	T	T	T	A	T	C	C	A	A	G	A	A	A	G	A	A	T	T	A	T	T	T	T	C	T	T	A	A	C	T	T	A	G	A	
3		T	T	T	T	T	C	T	G	G	C	T	T	T	A	G	A	A	A	A	A	T	G	C	T	T	G	T	T	T	T	T	T	C	T	T	C	A	A	T	G	G	A	A	A	A	T	A	G	G	C	A	
4		C	A	G	A	A	A	G	A	C	A	A	G	T	C	T	G	G	T	A	T	T	T	C	C	T	T	C	A	G	G	A	C	T	C	C	C	T	T	G	A	G	T	C	A	T	T	A	A	A	A		
5		C	A	T	C	C	A	T	C	T	A	G	C	T	T	T	G	A	T	T	T	T	T	C	C	T	C	T	T	C	T	T	C	T	T	T	G	T	A	A	G	A	C	T	G	A	A	T	A	G			
6		G	A	T	T	A	T	A	G	A	A	A	A	T	A	A	T	T	T	C	T	T	T	C	A	T	T	G	T	A	A	G	A	C	T	G	A	A	T	A	G	A	A	A	A	A	A	T	T	T	T	C	
7		A	T	A	C	T	T	T	G	T	T	A	G	T	C	T	C	T	G	T	G	C	C	T	C	T	A	T	G	T	G	C	C	A	T	G	A	G	G	A	A	A	T	T	T	G	A	C	T	A	C	T	
8		G	T	A	A	A	A	T	A	A	C	T	G	G	C	T	T	A	G	T	A	C	T	A	A	T	T	A	T	T	G	T	T	C	T	G	T	A	G	T	A	T	C	A	G	A	G	A	A	A	G	T	
9		T	C	T	T	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
10		T	T	T	A	T	G																																														
11		A	C	A	T	A	A																																														
12		T	G	C	C	T	G																																														
13		A	A	T	C	A	T																																														
14		T	A	A	A	T	A																																														
15		T	T	G	T	T	A																																														
16		T	A	A	A	A	A																																														
17		C	A	T	T	A	T																																														
18		T	A	A	G	C	A																																														
19		T	G	G	G	A	A																																														
20		T	T	T	T	A	C																																														
21		A	A	A	C	A	A																																														
22		A	T	C	T	A	C																																														
23		C	A	C	G	C	C																																														
24		T	C	G	T	C	G																																														
25		A	A	C	T	C																																															
26		T	T	A	T	C																																															
27		G	C	G	C	A	C																																														

.NET Bio Extension for Excel

### Select Input Sequences

Molecule Type: Auto Detect  Ignore blank cells  Treat blank cells as gaps

Query Sequence: gi343gbM125231HUMALBGCHuman1!N3:AH7

Name:

OK Cancel

Excel Biology Tool



# code talks

## let your voice be heard

CodePlex is Microsoft's free open source project hosting site. You can create projects to share with the world, collaborate with others on their projects, and download open source software.

[Terms of Use](#)

## Create Project

start a project

Join a project

[find projects seeking help](#)

Learn More

[faq and documentation](#)

Sign In

CodePlex Blog

[find out what's going on](#)

Improve Codeplex

[request features, log bugs](#)

Register

[register for an account](#)

MOST DOWNLOADED [1](#) [2](#) [3](#) [4](#) [5](#) [More](#)

### [Python Tools for Visual Studio](#) (67206)

PTVS is a free & OSS add-in from Microsoft that turns VS into a Python IDE. It supports CPython, IronPython, Intellisense, debugging, profiling, A...

### [Virtual Router - Wifi Hot Spot for Wi...](#) (33539)

The original open source Wifi Hotpot for Windows 7, Windows 8 and Windows Server 2012!

### [WBFS Manager](#) (26918)

POPULAR RELEASES [1](#) [2](#) [3](#) [4](#) [5](#)

### [FlowGraph](#)

**RELEASE:** [FlowGraph 1.0.0.0 source code - Apr 9, 2014](#)

The initial FlowGraph beta source code release.

### [NTLEA - NT Locale Emulator Advance](#)

**RELEASE:** [Ntleas Core.GUI Build 36 - Apr 7, 2014](#)

Version: 1.0.0b Build: 36 Description fixed some XUSE game crash problem, cuz VerQueryValueA

TOP FOLLOWED PROJECTS [1](#) [2](#) [3](#) [4](#) [5](#)

### [.NET Compiler Platform \("Roslyn"\)](#)

"Roslyn" .NET Compiler Platform

### [TypeScript](#)

TypeScript is a language for application-scale JavaScript development. TypeScript is a typed superset of JavaScript that compiles to plain JavaScript.

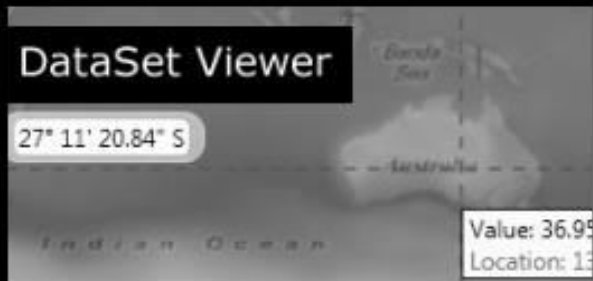
### [Visual F# Tools](#)

## Biocharts



Unify biological hypotheses with models and experiments

## DataSet Viewer



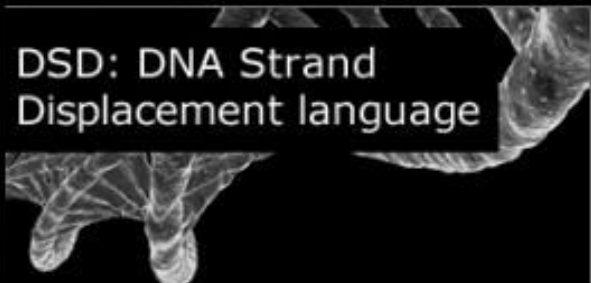
Easy visualization of scientific data: drag, drop, filter, slide, view, zoom, share

## Dmitrov

```
for (int i = 0; i < x.Length; i++)  
    xy += (x[i] - xm) * (y[i] - ym);  
double xx = 0; for (int i = 0; i < x.Length; i++)  
    xx += (x[i] - xm) * (x[i] - xm);  
var a = xy / xx; var b = ym - a * xm;  
var model = new double[x.Length];  
for (int i = 0; i < x.Length; i++)  
    model[i] = a * x[i] + b;
```

Easily work with multidimensional datasets: NetCDF, text, memory or remote, all from within your code

## DSD: DNA Strand Displacement language



A language for designing and simulating computational devices made of DNA

## Dynamic Data Display



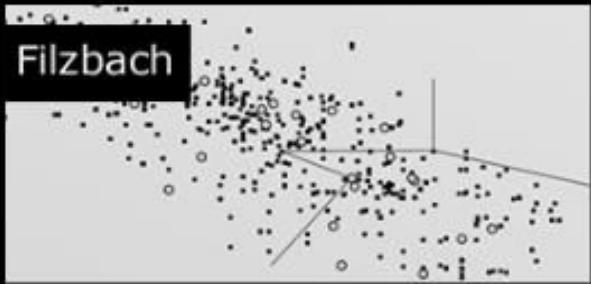
Visualize your data over the web and add complex dynamic graphs and maps to your web applications

## FetchClimate



Retrieve climatic and environmental information with the click of a button or a few lines of code

## Filzbach



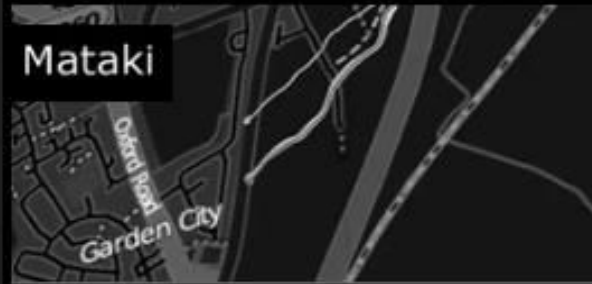
Fit complex models to heterogeneous data: Bayesian and likelihood analysis made easy

## GEC: Genetic Engineering of Cells language



A language for designing and simulating genetic devices to reprogram cell behaviour

## Mataki



Changing our understanding of animal behaviour: novel hardware, analysis and software tools

# Carlos Vásquez

Director Nacional de Tecnología y Competitividad  
Microsoft Colombia

[carlos.vasquez@microsoft.com](mailto:carlos.vasquez@microsoft.com)

---

# Gracias !!!!