Weapons lab to develop Celera's new supercomputer

ArticleInfo		
ArticleID	:	3966
ArticleDOI	:	10.1186/gb-spotlight-20010126-01
ArticleCitationID	:	spotlight-20010126-01
ArticleSequenceNumber	:	37
ArticleCategory	:	Research news
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2001–01–26 OnlineDate : 2001–01–26
ArticleCopyright	:	BioMed Central Ltd2001
ArticleGrants	:	
ArticleContext	:	130592211

Robert Walgate

Email: walgate@scienceanalysed.com

LONDON Craig Venter, the CEO of Celera Genomics - which is on the verge of publishing the sequence of the human genome - has signed an agreement with Sandia National Laboratory in the US to develop the most powerful computer in the world within four years - and it'll be used for biology.

At the same time, Sandia will be working on a similar machine to simulate the full three-dimensional impact of a nuclear weapons explosion, due for delivery to the US government by 2004. Sandia National Laboratory is a US Department of Energy laboratory, owned by the DoE but operated by the Sandia corporation. It is one of the three US weapons laboratories using supercomputers for the government's 'stockpile stewardship programme', which ensures the safety and reliability of the US nuclear stockpile in the absence of nuclear testing.

"The assembly of the human genome last June took 20,000 hours of central processing unit time" Venter said at the announcement of the agreement in Washington on 19 January. But to take the research further, "even though we have one of the biggest computers outside government," it had become clear that the biggest limitation would be computer power. And, said Venter, Celera has found the expertise it needs - in Sandia together with the personal computer company Compaq.

Compaq has already made a name for itself in biological supercomputing. It bought Digital Equipment Corporation - with its alpha chip technology - two years ago and provided the computers for the Celera human genome work.

Sandia's expertise is in designing "massively parallel systems" with thousands of computer chips working together, and the 'algorithms' for breaking down a problem so that a machine can solve it. Compaq has the chips, and growing experience in algorithm and supercomputer design and manufacture. The collaboration is estimated to assemble some 40-50% of the top algorithm scientists in the world. Celera's contribution to the project will be the biological problems, which will require new algorithms, and computer designs - especially involving large and continuous inputs and outputs to memory.

Last year, Compaq supplied several of the world's largest supercomputer systems, including the largest supercomputer in Europe to the French Atomic Energy Commission and the largest military supercomputer for simulating nuclear testing to the DoE. Accordiing to Compaq, Celera's machine for biology will be the largest supercomputer in the world.

Indeed, Compaq sees biology as the big future for the company. "We are already considered the world leader in supercomputers, and in bioinformatics," said a spokesman. "The needs are enormous and our people are extremely excited to be working on this."

According to Sandia's President, Paul Robinson, "we in the nuclear weapons community felt for many years that nothing could be more complex than nuclear physics and weapons... But I'm now convinced that nothing equals the complexity of bioscience and the human genome, and the challenges ahead. It's really important for us to have an opportunity to participate in this exciting work."

The director of the Celera project at Sandia, Grant Heffelfinger, echoes this. "It's tremendously exciting. It's the opportunity of a lifetime. It's like physics in the 1920s. Like the dawn of quantum mechanics. All of the fundamental questions are being challenged and assailed, we're just beginning to have some of the capacity to find the answers, and we find that the more you learn we see how much there is to know and how profound the answers are going to be."

The proposed machine so far has no name, but it will be "80 times larger than the current computer at Celera," Venter says. It will be "completely different" but complementary to IBM's Blue Gene project - which aims to predict protein folding from amino acid sequences.

Sydney Brenner, of the UK Medical Research Council's Molecular Genetics Unit in Cambridge, is sceptical of current bioinformatics, however. While admitting that "there are some problems where you can't do experiments," most post-genomic computing so far has been "a substitute for thinking."

In the December *Trends in Biochemical Sciences* Brenner writes that he had even heard it claimed that what he calls "-omic science" (statistical methods on genomes, proteomes and 'transcriptomes') "will liberate us from the domination of hypothesis, that is, thinking, in biology." But, he writes, many mRNAs and proteins may be useless, tolerated by evolution simply because they do no harm. "Only experiment can decide that."

Brenner told *BioMed central* that he does believe, however, in a future for what he calls 'computational biology', in which you might model the whole activity of all the components of a cell, down to molecular level. But he recognizes that we are a long way, so far, from having enough experimental knowledge to be able to construct such models.

Venter, it seems, has the same long-term target. The approach at Celera will be to integrate all biological information. "But it's going to another 10-20 years before we have a computer big enough to model how we go from a single egg and a sperm to the 100 trillion cells, 26,000 genes, 250,000 proteins that make a human body, with it's almost infinite number of combinations," he said. "So many interactions have to take place for one of our cells to be alive." But when this modelling is achieved "it will have a tremendous impact on our understanding of cancer or pharmaceutical development. We think that we can save decades in the development process" said Venter. "We need this collaboration to go from genome mapping to biology."

In terms of 'floating point operations per second', or FLOPS, the new machine will aim first at 100 TeraFLOPS, and then 1000 TeraFLOPs, or 1 PetaFLOP (1015 FLOPS) by simple scale-up. The largest machine in the world, at Lawrence Livermore Laboratory in the US, reaches 12 TeraFLOPS. Blue Gene's target is also 1 PetaFLOP, but its algorithms and architecture are likely to be different from that of Celera's machine, as the machines will be designed to solve different problems. Some believe that by virtue of aiming at all post-genomic biology the Celera machine will prove to be the more versatile, and will be applicable to national security issues including rapid message encryption and decoding - a key component of spying.

In reply to concerns about the possible links with defence and security in this project, a Celera spokesperson said: "It seems that once again there is this notion that there must be something sinister and negative about this research agreement. The only way that we are ever going to be able to better understand how the human body functions is by creating more powerful computing tools, and if this

collaboration can enable those advances I should think that would be a very good thing for all of us. I can't think of an 'X-Files-like' scenario in which it would benefit anyone to 'hive off' the information. Don't forget we are publishing our human genome information and posting that information on our web site; and there is a public consortium posting data to their web sites. I know those science fiction depictions are fascinating but most times they are just that... fiction."

References

1. Celera Genomics, [http://www.celera.com]

2. Sandia National Laboratories, [http://www.sandia.gov]

3. US Department of Energy press release on the supercomputer agreement, [http://www.energy.gov/HQPress/releases01/janpr/pr01022.htm]

4. Compaq Computer Corporation, [http://www.compaq.com]

5. IBM's Blue Gene, [http://www.research.ibm.com/news/detail/architecture_fact.html]

6. US Department of Energy security pages, [http://www.energy.gov/security/index.html]