TIP White Paper

CRITICAL NATIONAL NEED AREA

A National Infrastructure Standard for 21st Century Genetics

"The DNA-Net"

Submitting Organization

GenVault Corporation

6190 Corte del Cedro

Carlsbad CA 92011

www.genvault.com

Contact name

Michael Hogan, Ph.D.

Chief Scientific Officer

GenVault Corporation

760-268-5200 (office)

520-904-1715 (mobile)

mikehogan12751@earthlink.net

Keywords:

Networks, Applied Genetics, Biotechnology, Biobanking, Personalized Medicine, Diagnostics, Forensics, Environmental testing

Executive Summary.

Human genetics is advancing at an exponential rate, towards a world where many aspects of commerce, human health and national security will be managed, based on data obtained from complex genetic screening. Below, we highlight how far that reality has progressed as of today, with a look to 2015, where the role of genetic testing in health, security, commerce and the environment will have grown to @100 million tests/yr.

Table 1. How Genetics will become a Central Feature of U.S. Society in 2015

Application Area	2009 Status	2015 Status (proposed)
Medical Treatment-Transplantation	solid organ & marrow	solid organ & marrow ***100K tests/yr***
Medical Treatment-Stem Cells	R&D only	Neuro-degenerative, CV, plastic surgery: genetics to determine a match ***100K tests/yr***
Medical Treatment-Pharmaceutics	warfarin, abacavir	*Nearly all drugs* based on liver clearance, immunological rash, receptors ***1MM tests/yr***
Medical Treatment-Vaccination	R&D only	universal HLA screening for childhood and adult vaccine response ***10MM tests/yr***
Medical Treatment -Diagnostics	BRACA, EGFR (cancer)	All cancers, all CV indications, rheumatology ***1MM tests/yr***
Neonatal Screening	CF, cycle cell, Tay Sachs	HLA and at least 5 others, universally at birth. This will be the 21st century analogue of the ABO blood type ***1MM tests/yr***
Public Health-Infection Risk	MRSA, classical petrie dishes analytical culture for the rest	HIV, flu, dengue, West Nile sensitivity, MRSA, drug resistant TB risk ***1MM tests/yr***
Public Health-Cancer Risk	R&D	Liver genes (carcinogen clearance) Immune markers, DNA repair ***1MM tests/yr***
Public Health-Obesity Risk	R&D	Ghrelin, adipocyte stem cells, metabolism, neuro-addictive screening, ***1MM tests/yr***
Environmental Testing-Water	Classical petrie dishes & culture	Genetic testing of 40,000 sources per day in U.S. will replace cell culture, ***10MM tests/yr***
Environmental Testing-Air	Classical petrie dishes & culture	Flu, drug resistant TB, will replace cell culture, ***1MM tests/yr***
Environmental Testing-Soil	Classical petrie dishes & culture	Microbial biodiversity via microarrays, to monitor contamination, ***100K tests/yr***
Climate Change-Biodiversity	Smthsonian & NSF bar code of life project	Worldwide genetic biodiversity testing, especially among sentinel microbes, ***1MM tests/yr***
Food Testing-food borne disease	Some PCR but classical petrie dishes & culture dominate	Microbial screening for all domestic and imported food stock: factories & point of entry via PCR, beads, microarray, 5***0MM tests/yr***
Food Testing-genetic modification	Genetically engineered plants: QC and industrial piracy	Large scale screening of environmental back crossing to wild stock, ***10MM tests/yr***
Forensics-casework	All ID via Identifiler	ID via Identifiler & trait & clan analysis, ***10MM tests/yr***
Forensics-ID databases	All violent offenders	All booked offenders,***1MM tests/yr***
Defense-Military	Identification of the dead	Identification of the dead, performance trait screening of all recruits, ***100K tests/yr***.
Defense-Bioterrorism	R&D	Bioshield air and water screening for militarized pathogens, ***1MM tests/yr***
Defense-Immigration	R&D	Current French model: DNA ID on all visa applicants, ***1MM tests/yr***
R&D-Human Genetics	Traditional lab R&D on small ad hoc biobanks	National discovery biobanking similar to UK, Spain, France, Canada, Lux, Singapore, Japan, Malasia, Australia, ***1MM tests/yr***
R&D-Animal Breeding	Chicken & bovine quantitative trait selection at DNA level	Expansion to all feedstock for marker based selection, ***10MM tests/yr***
R&D-Plant Breeding	Corn feedstock	Corn, soybean, wheat, rice and fuel stock for food and biofuels: ***10MM tests/yr***

Applied Genetics: Woven into the Fabric of Society. The range of genetic knowledge, although already impressive in 2009, is predicted to double on a yearly basis over the next few decades. Thus, it is clear that we have entered an era where genetic principles and genetic testing to back them up, *will be woven deeply into the fabric of U.S. healthcare, commerce, safety and environmental protection.* The field is early enough in its development, that it is still not clear what the full range of genetic testing will be, or over time, what the spectrum of technologies will be to support the full range of genetic testing that will emerge over the next fifty years.

Guiding Principles. In spite of that technical uncertainty, we believe that four general principles can be laid out, today, to guide the future of ***societal-scale*** genetic testing, as in Table 1, over the next 50 years. We believe that these principles can be used by the U.S. to become the international leader and, consequently, to reap the leader's reward from defining the way that 21st century genetics will be built into the fabric of world society.

1). Complexity. The available science has shown that genetic factors which lay beneath the diverse and important applications of Table 1 will not usually be revealed by simple single-gene tests, but will involve analysis of relatively complex gene panels, often at the complex allele level, rather than at the level of simple localized gene polymorphism.

2). Strategic Planning. The genetic testing that pervades U.S. society in the mid 21st century will be correlated with well-defined risk factors: specific inherited birth defects, specific medical treatments, identified risk of exposure to a chemical or biological agents, standardized forensic markers, standardized screens for viral and bacterial exposure, and so on. Thus, high value genetic testing will not ordinarily be done in haste, in the field or in the home, in minutes (the so-called point of care or "tricorder" approach, that is much talked about) but will in most cases be done diligently, at birth, or in preparation for a medical treatment, or in preparation for military recruitment, or during standardized programs of food, or air, or water quality sampling.

3). Centralized Testing. The technologies that will enable such large scale, genetic testing will be very high throughput, and multiplex in nature, thus minimizing the amount of DNA that must be collected per individual or food or environmental sample. That kind of very high throughput, multiplexed analysis will almost certainly be performed at a few specialized sites, which, generally speaking, will be at great distance from any particular site of medical, or forensic, or industrial, or environmental sampling. The results of such testing, as complex data, will then be routed to yet other sites, where they can be re-compiled and used for complex decision making.

4). Networking. If, as hypothesized in Table 1, complex genetic testing becomes a routine component of commerce, defense, healthcare, public health & environmental screening in 2015, we will have undergone a transition from 2009, where almost no one has benefited directly or indirectly from any genetic test, to a situation in 2015, where genetic analysis has become as routine in everyday life as applying for a bank card on the internet. Like the bank card analogy, the practical value of genetic testing in 2015 will involve coordinated flow of information from the individual, through a network, where the information is analyzed centrally, so that the product of that analysis may be used as information to guide everyday life, for individuals, companies & states. Thus, we imagine that, like the bankcard, the world of 2015, having been permeated by complex genetic testing, must, correspondingly, be supported by a genetic network with formal similarity to the pervasiveness of the Internet, as we have come to know it, over the past 20 years.

Based on consideration of those 4 principles, we perceive a fundamental technological weakness, for consideration here by NIST-TIP, which, if left un-remedied, could block the development of such an extended genetic data network and would, as a result, greatly limit the societal value of genetic testing, for decades.

The Genetic Testing "Disconnect". At its heart, complex genetic testing is an example of sophisticated physical chemistry (the physical chemistry of the DNA polynucleotide strands) coupled to sophisticated informatics, which is used to assemble sequencing chemistry or to de-convolute hybridization binding interactions into gene sequence structure. In spite of its direct coupling to such "21st century polymer chemistry", the acquisition, transport, storage and purification of DNA is, for the most part, still treated like an exercise in functional biology. DNA-containing samples are treated as if they were "alive", rather than polymer chains: they are shipped in the cold, stored in the cold and subjected to purifying treatments that were developed in a 20th century world of "wet" bench biology, rather than with an eye to supporting an extremely high-tech marriage of

physical chemistry and computer science. As a result, high throughput, computer intensive, applied genetic analysis--the future of societal-scale genetics--has become captive to, and ultimately bogged down by, the slow, expensive, arcane methods of 20th century DNA sample collection, preservation, shipping & recovery.

An Infrastructure to Enable Societal Genetic Testing for the rest of the 21st Century.

Historical Perspective-Solid State DNA. Although not generally recognized as such, the pervasive role that genetics will play in the U.S. economy, medical care and security, in the later half of the 21st century, began rather humbly 50 years ago in 1962 (before gene structure was even known) in an application area that was only indirectly related to genetic testing. At that time, Dr. Robert Guthrie, one of the U.S. pioneers in neonatalogy, became concerned that ordinary medical sample collection would be too expensive to support the first generation of birth defects testing that had just become available, to test for the failure to process phenylalanine (PKU).

Although birth defects are usually inherited, and thus ultimately DNA-based, the PKU test assessed phenylalanine levels directly in newborn blood. Since it was a simple small-molecule that was to be tested, Guthrie wondered if a standard (expensive) medical blood draw could be eliminated from neonatal screening, and instead the blood would presented as droplet on the skin, via a heal stick, to be soaked directly into filter paper. After that, the blood spot could be air-dried and stored ***in the solid-state*** for processing at a regional site, much like a postcard is sent through the mail. Guthrie's brilliant insight turned out to be correct, and as a consequence, simple blood-spot transfer to filter paper in a "Guthrie Card" has remained the neonatal screening standard 50 years, with no technological change at all. In 2009, the simple filter-paper Guthrie card enables neonatal screening on about 60,000 babies per day in the U.S., E.U. & Japan, even though the small molecules in such "lowly" Guthrie cards are analyzed via sophisticated, automated, tandem mass spectrometry.

With the invention of PCR, about 10 years ago genetic testing (slowly) began to enter neonatal screening, specifically targeted to rare, relatively-simple genetic traits such as cystic fibrosis, sickle cell anemia, Tay Sachs, and a few others for which one or a few PCR reactions would suffice. In spite of that transition from small molecule to small molecule plus targeted genetic screening, the simple solid-state Guthrie Card approach has been shown to present enough blood DNA, in good-enough quality, to accommodate both complex molecular testing and a few simple genetic tests from a single dried neonatal blood spot.

Extrapolating the Solid-State Guthrie Card into the Future. Going into the mid 21st century, Guthrie's basic wisdom of collecting, transporting, storing and processing DNA in the solid-state becomes even more important. We are entering an era where genetic testing will be 10-100 times more complex than simple single gene tests such as cystic fibrosis. We are personally aware of a single agricultural company that performs 100,000 PCR-based genetic tests per day, at one lab site, on corn only. The State of California plans to perform 250,000 genetic ID tests on accused offenders in 2009. Our best estimate is that among public health, medical treatment, neonatal screening, forensics, military screening, agricultural testing, food quality testing, and environmental monitoring (as in Table 1) by 2015, genetic analysis will be performed in the U.S. at a rate of about ***100 million tests per year***. Nearly all of those tests will be performed at specialized regional test sites. To envision the enormity of the problem that awaits, and the nearly complete inadequacy of the existing genetic management infrastructure, we propose the following representative scenario, based a single high-likelihood example, in which genetic testing will become standard part of public health vaccination (Table 1).

Model: Genetic Testing for Vaccination, in 2015. Imagine it is 2015. Studies begun at Mayo and other top labs in the 1990's now show that natural variation in the immune system (the same variation that gives rise to transplantation rejection) also gives rise to significant personal variation in the effectiveness of standard vaccines. Thus, as part of the ordinary vaccination process (10 million vaccinations per year in the U.S.) a parent has agreed to have their child tested for a complete panel of HLA gene markers, to teach which vaccines will work, and those that will not, or that will be associated with side-reaction on the child. The analysis is performed by fourth generation re-sequencing or microarray technology, which covers about 1mB of the genome (for about $100) and will consume a total of about 1ug of total DNA, which is readily obtained from a drop (50uL) of human blood from a healthy child or adult volunteer. To get the test costs down to $100, such public-health-scale HLA testing will be performed at a secure, regionalized, very high throughput genetic testing facility, in another town or state.

The Fundamental Problem of Scale is Revealed. HLA testing is nicely representative of the future of complex multiple gene analysis. It is done every day in the U.S. in support of organ transplantation, at about 100,000 transplantations per year. Transplantation medicine is well known to all, and as of 2009, it has become a central feature of medical practice. However, on the scale of societal genetics in 2015, (100 million tests) transplantation genetics would be viewed as a "niche" application. In 2009, the sample collection standard for HLA genetics is an ordinary venous blood draw, which in the context of a transplant, is easily collected in a hospital, for a cost of about $200—a small price to pay relative to the $20,000 to be spent on the transplant itself.

When, as widely envisioned, HLA typing becomes the coupled to the delivery of public health vaccination, at 10 million vaccination decisions per year, rather than 100 thousand transplantation decisions per year, that 100 fold ***increase*** in market size and the 200-fold ***decrease*** in the value of the corresponding vaccination treatment, makes it easy to see that a costly venous blood draw, followed by refrigerated shipping of the blood to a regional center, cannot be justified in terms of societal value & price. Moreover, cost aside, the logistics of routing millions of refrigerated blood tube specimens from the site of collection to the site of use, creates an organizational bottleneck that simply cannot be sustained in the scale of societal testing, envisioned in Table 1.

Thankfully, there is a Genetic Version of Moore's Law. Operational costs aside, going into the mid 21st century, there is simply no justification for acquiring physically large samples for genetic analysis. Much like Moore's law for integrated microdevices, the amount of DNA specimen required to perform complex genetic testing has been dropping by more than a factor of ten, per decade. For instance, in 2000, re-sequencing a 1mB patch of a genome was the activity of a small research laboratory, requiring about six months, about a $100 thousand, and about 100 micrograms of DNA. With third generation re-sequencing, it can be now done in 2-3 days, for about $1,000 and about 2ugs of DNA input. Similarly, in 2000, single nucleotide polymorphism, done as one PCR reaction at a time would require 1,000,000 PCR reactions at $1 each, over a year, with about 1000ug of DNA, in order to scan the human genome for genetic variation at high resolution. In 2008, a standard Affymetrix or Illumina microarray will do the same task for about $500 dollars, in a day, with about ½ microgram of DNA input. That is far less than amount of DNA that can be obtained from a single drop of blood. By 2015, the cost, time and sample requirements will almost certainly be reduced by another factor of 5-10. Thus, in the mid 21st century and beyond, a drop of blood (which contains about 200,000 copies of the human genome) becomes a very large, not a small sample in the context of complex genetic analysis.

21st Century Solution to Routing 100 Million Specimens a Year. Rather than ordinary fluid sample collection, we propose that for HLA and most other complex testing applications to come over the next decades, ***transition of the DNA sample into the solid state will be enabling.*** Imagine that the blood drop to be used for HLA typing is presented by a painless finger prick (rather than with a syringe) obtained while standing in line and transferred to a "Collector", essentially the 2015 version of the Guthrie Card, except that it cannot be a card at all, since a thin paper Guthrie Card is relatively flimsy, hard to ship, or store in a biobank for routing; is unprotected, cannot be used to purify DNA; has limited storage capacity; is hard to automate; and presents the specimen with open access to the air around it: where the specimen can be contaminated, or in turn, can contaminate its surroundings. Still, the formal similarity between the Guthrie Card and what will be needed in 2015 is strong enough to justify describing that future, hypothetical, dry-state specimen technology, the "Collector", as the 21st century grandson of a 1960 Guthrie Card.

We envision that the "Collector" might look like a USB jump drive. The Collector would be touched to the finger; the blood drop is transferred directly into it by passive wicking; the collector would then be immediately re-capped; and the embedded rf tag read, to enter the specimen into a network wherein it can be shipped a great distance, banked, routed and re-distributed. At a processing site, the rf tag on the Collector is re-read, to confirm source, identity and use; the Collector is opened and soaked for 20 minutes at room temperature to rehydrate the DNA specimen, which is then washed to remove protein contaminants and the DNA eluted in pure form into solution for analysis.

All of the processing steps just described would be performed with ordinary (2009) laboratory automation, and, most importantly, the DNA thus obtained would have been drawn from the child, or adult, stabilized, shipped, stored and recovered in the same Collector standard. At no time would the specimen be refrigerated and, until the point at which the purified DNA is finally used for genetic analysis, the DNA would have resided in the same Collector throughout acquisition, shipping, storage, re-hydration and purification.

There is nothing in our current understanding of biotechnology that would provide such a simplified "Collector" workflow, and so, in order to accommodate multiple-gene testing on many millions of individuals and environmental samples per year, fundamental (perhaps revolutionary) enhancements to the 50-year old Guthrie card idea will need to be invented.

This new technology must account for the flow of genetic material from a very large number of collection sites (sources) then routing to multiple specialized sites of analysis (receivers). When described in that way, it can be seen that societal-scale genetic testing is, in fact, a **network-based** problem; one that is at least 1000 times more complex than our current understanding of the logistics of universal neonatal screening. We propose that very large-scale, ***societal genetic testing*** will become formatted as a network, with properties similar to the flow of electronic information, as embodied in our current understanding of the Internet, and the way that the Internet is coupled to complex physical routing systems such as FEDEX or USPS.

For the purposes of discussion, we use the term **"DNA-net"**, to refer to the network solution that will be required to collect, route and distribute physical genetic information in the mid-to-late 21st century. It will be a national or international scale system to orchestrate the flow of physical genetic content, embodied as DNA strands in the solid state. With only minor technical and formatting modification, the same DNA-net would support the diverse applications of Table 1, as a secured network with access and interoperability that would be regulated by the same sort of password protection, encryption and firewalls that have been developed, very successfully, for the Internet.

What would a DNA-net look like as a complex structure? The DNA-net will have an interesting formal analogy to other complex networks that had been developed to move information, both LAN and the Internet, and would control the complex flow of genetic material, much as the internet-enabled FEDEX or USPS model controls the complex flow of physical packages. Below, we use the standard descriptive formalism of the Internet to describe ***the 3 components that would need to be invented*** to create such a very large scale DNA-net.

A). Transduction of Content into a Standard

In the Internet, the underlying format is based on the transduction of diverse information types (the content) into a standardized binary code. ***In the DNA-net***, the physical formatting standard would be transduction of DNA in diverse sample types (the content) into a standardized solid-state format. Much is already known about maintaining DNA in the solid state for genetic analysis, and that current knowledge would be expanded upon to create the new solid-state genetic content standard for the DNA-net.

B). Formatting of Content for Transport & Tracking

In the Internet, all binary code content is formatted into a standardized "Packet", where the binary content is parsed into a standard size (the "Payload") and wrapped with a "Header", which includes important information about the Payload. The Header has a standard information format which includes a description of the type of content that is in the Payload, where it has come from, and where it must go. ***In the DNA-net***, the "Carrier", as we have defined it, would function as the "Packet"**.** The solid-state DNA content would be parsed into a uniform Payload format and size (a solid DNA alloquot of a standard dimension) and the Header function would be embodied in the rf-tag. As for the Header of a standard TCP/IP internet Packet, the rf tag would identify the type of content in the "Carrier" (DNA from blood, DNA from saliva, DNA from plants, DNA from water filtrate, etc); where it came from (a hospital, a police station, a water treatment plant); and where it must go (a centralized medical testing facility, a crime lab, a water analysis lab). The ability to format data as a Packet, is arguably, the underlying core technology of the Internet. Similarly, the invention and optimization of such a "Carrier", will become the enabling core technology of the DNA-net to come.

C). The Router.

In the Internet, Packets are shipped throughout a network linked by nodes, where each node operates as a Router; that accepts Packets that have been delivered to the node; and then routes them to the destination address specified on the Header, in the fastest way possible. In the Internet, there are two functionally distinct types of Router: **a).** Routers that manage the flow of Packets within a local network and **b).** Routers that manage the flow of Packets between local networks. Generally, the Local and Network Routers are owned by different institutions: your company or university may own the Local Router, which operates on a cable network system that is also owned by your company; while the Network Router may be owned by a Service Provider (ISP) such as Earthlink, who use a physical network that is based on fiber optics or satellites, which may, in turn, be owned by a third party such as a phone company. ***In the DNA-net***, the Local Router will be a new type of standardized automated system, not yet invented, to manage, store and retrieve Carriers on demand; while the Network Router would be provided by the FEDEX or USPS system as we know them, who would route Carriers from a Source to a designated Receiver, via the existing physical network system (highways, air-routes) that are owned by third party, usually the Federal Government.

How long would it take to create the DNA-net, and how much would it cost??

Since the Network Router function (FEDEX, USPS) and physical network infrastructure (highways, air-routes) already exist for the proposed DNA-net, the rate limit to the development of the DNA-net will be development of the "Carrier" standard (i.e. the physical analogue of the Packet) and the corresponding "Local Router" function, a type of high throughput robot to manage the storage and redistribution (the routing) of the Carrier. It is our best estimate that with NIST-TIP support, by leveraging off-the-shelf components & technologies, the needed Carrier and Local Router function could be delivered in a fully operational form ***in 36 months***, with full compatibility to the existing Internet overlay and the existing FEDEX/USPS Network Router infrastructure, for a cost of ***less than $10M***. The deliverable would be a fully operational prototype, comprising two local networks (each at a University or National Lab) each with its own Local Router, compatible with both the FEDEX and USPS Network Routing format.

It is our best estimate that, if the TIP program made the DNA-net Program a funding priority in 2009, up to 10 very high profile academic, industrial and academic-industrial hybrid teams would stand ready to complete for the opportunity to design and build a prototype solution. A full system, to accommodate 50 years of growth in the DNA-net, in support of the full range of DNA-bearing sample, could be completed in an additional 5 years, we estimate, for less than $50M.

What would be the immediate value of the DNA-net??

I. Enhance short term scientific and IP value: catch up in the biobank race. The U.S. has been responsible for nearly all of the knowledge and technology that underlies the current genetic understanding of disease origin, disease risk, response to medication, forensic genetics & the genetics of food quality & microbial diversity; virtually all of the "2009" column of Table 1 had its origins in U.S. science. However, the next step in the evolution of such technology and the related knowledge base (the 2015 column of Table 1) will be based on the ability to perform complex genetic discovery analysis in the context of very large & secure networks of specimens, that are coupled to medical, or public health, or military data (referred to, these days, as biobanks). The U.S. has fallen behind the E.U., Canada and Asia, ***by at least 5 years***, in the development and use of such biobanks to discover new, high-value, genetic correlations. We continue to drop farther behind the foreign competition, for two reasons:

1). Socialized Healthcare, Agriculture & Food Safety. The first reason for the compromised U.S. competitive position, often sited, is that foreign healthcare systems are much more centralized, thus making it easier to obtain ethical approval for large-scale human discovery research, and, in the area of agriculture and food safety, the socialism makes it easier to support such industrial activities, from taxes.

2). An Interest to Invest in Applied Science. The second reason, less often cited, is that over the last 8 years, the Canadian, E.U., Australian & Asian governments have invested heavily in biobanking networks, with the clear intention to use those biobanks as a genetic discovery engine. Meanwhile, there has been essentially no U.S. investment in biobank network infrastructure development, or direct funding for such national-scale biobanks. As

a result of those early, national-level investments in genetic "content" generation, it is likely that many of the patentable discoveries that relate genetic variation to disease, to agricultural vigor, or readiness to military attack (Table 1, column three) will be owned by foreign nations, even though most of the basic knowledge and technology used to support those practical discoveries was invented here.

II. Attain long-term commercial & defense superiority in applied genetics. The U.S. must come to grips with the idea that genetics will become woven, intimately, into the fabric of 21st century society, and that the core of that genetic value will manifest as a large-scale network. First, it will manifest as local genetic networks (biobanks) to discover high-value genetic correlations in health, commerce and security. It will then manifest as larger genetic networks, where that research knowledge is put to work, to acquire, compile, route and use large quantities of genetic material & associated information. It is certain that, over the next 10 years, the U.S. healthcare system cannot be changed fast enough to make the U.S. competitive with socialized economies in the E.U. and Asia. However we argue strongly that, ***in much less than 10 years***, the U.S. can become very competitive by doing what it does best, which is to innovate technically: thereby leapfrogging the Canadians, the Europeans and Asian competitors, by funding the invention & commercialization of a DNA-net infrastructure which, ***in the short-term***, will balance-out the early biobank IP discovery lead afforded to the competition; then ***over the long haul***, allow the U.S. to regain leadership in the delivery of applied genetics knowledge, as the newest high-value addition to the fabric of U.S. research, commerce and security.

III. Marching orders. A useful analogy: early federal support of the Internet. *If* in the 70s & 80s, European & Asian computer science were 5-10 years ahead of that in the U.S. and *if*, also in the 70s & 80s, DARPA and then the NSF had not chosen to fund infrastructure building, for what became the Internet, *then* the Internet-based society we live in now, would not have the distinctly U.S. look and feel to it, that it does.

Instead, the Internet of 2009 would be much more like the automobile industry of 2009: an industry that has become completely dominated by imports. In 1962, when Guthrie stumbled on the concept of solid-state biospecimen storage and the "routing" of such solid state biospecimens for universal neonatal screening, other than a few visionaries, most scientists had no idea how the next 50 years of genetics would change U.S. commerce and society. Also in 1962, as it turns out, when Licklider at MIT first conceived of his "Galactic Network" of linked military computers, the seminal ideas that foresaw Internet, no one except the greatest visionaries (perhaps not even Licklider) could have anticipated how the Internet would change society in the last part of the 20st century. From the perspective of 2009, it no longer takes a visionary to imagine that, over the next 50 years, genetics, particularly the flow of genetic material and genetic data, will change the world in ways that will begin to match how networks of pure electronic data changed the 20th century.

The challenge, we believe, is to recognize that genetics in the 21st century will no longer be an exercise in wet laboratory science; no more so, than the Internet of today is an exercise in the material science of tube based amplifiers, copper wires or contact switching. As is the case for the materials science beneath network informatics, the role of genetics in society is quickly evolving into a type of complex network application. Here, we have called it the DNA-net, for short, where the value to be had from genetics will be guided by, and possibly enabled by, new physical networking technologies (yet to be invented) that will manage the flow of, and controlled access to, those high value genetic assets, and the data associated with them.

DARPA (in 1972) and the NSF (in 1988) took a visionary leadership position, to make the Internet a reality in the 20th century. We argue that, in 2009, ***NIST is in a position to assume that same kind of leadership***: to guide, via support of the underlying infrastructure, the way that networked genetics will permeate 21st century society. Although the investment made by DARPA and NSF was ultimately commercialized and made available to the world, the early U.S. leadership role that they took, subsequently propagated into many commercial Internet spin-offs that allowed the U.S. to retain commercial superiority for decades thereafter. The same argument can be made for the DNA-net. If NIST and others in the U.S. choose to take a leadership role, to build a DNA-net, the resulting technology will also be commercialized, and sold internationally. But if managed correctly, early leadership in defining and developing DNA-net standards for the 21st century will also propagate into U.S. commercial superiority in applied genetics, for decades to come.