

# GAO ANALYSIS HIGHLIGHTS LAB SAMPLES EXCLUDED IN SLOPPY FBI ANTHRAX INVESTIGATION

As the last Friday before Christmas, late yesterday afternoon was the most obvious Friday news dump hour of the year, and the government didn't disappoint. The Government Accountability Office released the results of a twenty-three month long study of the genetic analysis that was used to tie the material found in the anthrax attacks of 2001 to the laboratory of Bruce Ivins, whom the FBI concluded (pdf) was solely responsible for the attacks. The FBI's conclusion is highly suspect for many reasons. On the science side, it is very unlikely that Ivins could have produced all of the attack material on his own and the detailed chemistry of the attack spores suggests that highly sophisticated materials and techniques unavailable to Ivins likely were used to prepare the attack material. Regarding that second point, note that even William Broad refers indirectly to the chemistry concerns in his New York Times article on the GAO report:

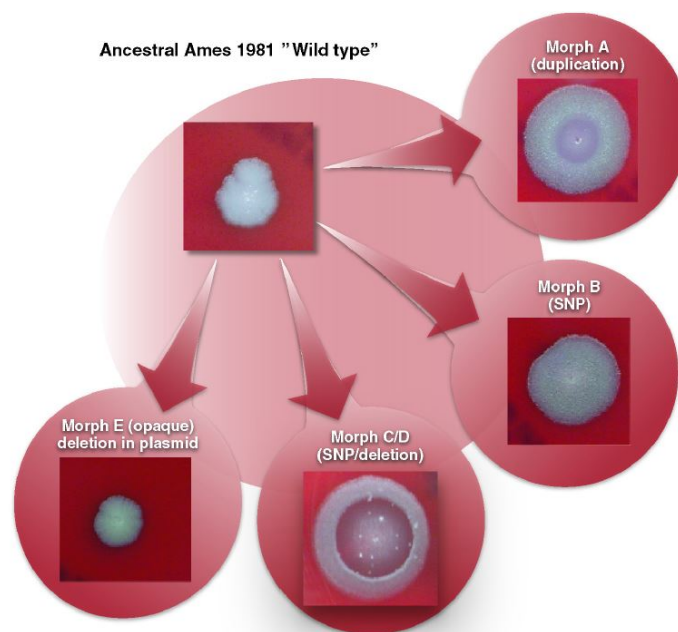
To the regret of independent scientists, the report made no mention of an issue beyond genetics: whether the spores displayed signs of advanced manufacturing. They have pointed to distinctive chemicals found in the dried anthrax spores that they say contradict F.B.I. claims that the germs were unsophisticated.

Evidence of special coatings, they say, suggests that Dr. Ivins had help in obtaining his germ weapons or was innocent.

The GAO study was undertaken, in part, because of questions raised by the National Academies study released in 2011 and with special prompting by Representative Rush Holt, from whose district the letters likely were mailed. The GAO study focused on obtaining a better understanding of the validity of the genetic analysis that was carried out and the statistics underlying the conclusions reached.

For a refresher, a helpful illustration from the GAO report shows the underlying biology of the genetic analysis that was carried out in the Amerithrax investigation. Here we see photos of a typical colony of the Ames strain of *Bacillus anthracis* on an agar plate and four variant colony types that occurred at low frequency when the attack material was spread out on agar so that colonies arose from single cells of the overall population of bacteria that were present in the attack material:

Figure 2: Ancestral Ames Strain and Types of Morphs Found in the Evidence from the 2001 Anthrax Attack



Sources: GAO and photographs courtesy of USAMRIID. | GAO-15-80

DNA sequence analysis was employed to identify the changes that led to these variant colony shapes. The FBI then commissioned private laboratories to develop DNA-based tests (relying on polymerase chain reaction, or PCR, methodology) that could be used to screen the large bank of isolates of the Ames strain that

the FBI had accumulated through a subpoena submitted to all 20 laboratories known to have isolates of the Ames strain. Developing these assays represented a new frontier in forensic genetics and it did not prove possible to develop tests for all of the mutations identified in the original DNA sequencing. In the end, four tests were developed by the four different contractors.

The Amerithrax report stated that of the 947 samples included in the final analysis, only eight showed all four of the DNA changes the tests were designed to detect. Seven of those samples came from the laboratory where Ivins worked (U.S. Army Medical Research Institute of Infectious Diseases, or USAMRIID) and one came from Battelle Memorial Institute in Columbus, Ohio. The FBI noted that there was a record of material being transferred from USAMRIID to Battelle, accounting for the sample found there.

The GAO analysis finds a number of significant issues with the FBI's work:

#### **Source of Variant Types**

First, the GAO report noted that during the development of the genetic tests, questions arose about the factors underlying the presence of variants and especially whether culture conditions might affect the relative populations of normal and variant types:

Although the specific genetic mutations used as genetic markers to determine a match or exclusion were adequately characterized, the FBI did not conduct studies to understand the methods and environmental conditions that gave rise to the mutations. The FBI convened a team of scientists in 2007 to review the scientific methods. Finding no shortfalls or deficiencies in the basic methodologies they reviewed, they determined that the usefulness of the genetic markers was sufficient. The team also stated that the extent of research

and development of the genetic tests at the date of their review was insufficient to determine whether the presence or absence of one or several of the genetic markers was associated with the evidence, was merely characteristic of normal culture practices, or possibly was affected by the sensitivity of detections of the genetic tests. The team recommended additional studies to characterize the genetic markers as a function of growth conditions, including the influence of growth time, growth media, and temperature.

The GAO reports that the FBI's response to these concerns when they were raised by the NAS panel was hardly encouraging:

In response to questions from the NAS panel about this recommendation, the FBI stated that it considered such studies academic and did not conduct the recommended research.

But that is hardly a just an "academic" question. See this post of mine for a summary of the preparation of Ivins' RMR-1029 flask, which the FBI treated as essentially a smoking gun. That flask had material from a large number of large scale cultures. Also, the sheer amount of very highly concentrated material in the recovered letters from the attack also suggest very large cultures were carried out to produce the attack material. By comparison, the material submitted by the laboratories in response to the subpoena would be from very small laboratory scale cultures, and so the growth conditions would have been quite different, quite likely affecting the ratios of variant types in the final populations produced.

### **Sample Submission**

Besides the concerns about culture conditions affecting the presence of variants in the

samples submitted, the NAS report highlighted a point that had been somewhat obscured previously. It turns out that the scientists responding to the subpoena showed huge variations in how they responded and what they considered to be separate laboratory populations worthy of sample submission:

Our analysis of FBI documents shows that FBI searches at three specific laboratories identified hundreds of additional relevant stocks that laboratories did not submit to the repository in response to the subpoena. Specifically, we found that the FBI collected about 29 percent of the 1,059 repository samples through these searches.

That's staggering. Nearly a third of the total repository of samples would not have been present had the FBI not searched those three labs. From the Amerithrax report, we do learn that the three that were searched were USAMRIID, Dugway and Batelle. But what about the 17 sites submitting samples that weren't searched? How many populations were missed in the pool that was tested? The bottom line is that the FBI analyzed a pool of samples that very likely missed a huge portion of what should have been analyzed.

### **Validation**

Very far into the process of developing the DNA tests, the FBI realized they needed to make an effort at validating their analysis. One of the validation attempts put one of the tests into huge question. Table 3 from their report shows this disappointing result:

Validation testing showed that for those results expected to be positive, no negative results were observed at or above the LOD for any of the genetic tests.<sup>40</sup> However, in the postvalidation testing, the negative rates were generally high. As shown in table 3, the negative rates for the postvalidation tests ranged from 0 percent to 43 percent for the undiluted samples from flask RMR-1029. (Appendix III breaks down the results of the replicate testing for each genetic test.)

Table 3: Sensitivity Results for Five Postvalidation Tests on Undiluted Samples from Flask RMR-1029

Genetic test	Number		Sensitivity	
	Replications from flask (positive samples)	Positive samples detected	Nonpositive results <sup>a</sup>	Estimated % negative rate <sup>a</sup>
A1	30	17	13	43.3
A3	30	29	1	3.3
D-1	30	23	7	23.3
D-2	30	24	6	20.0
E	30	30	0	0

Source: FBI, sensitivity statistics derived from 30 replicate samples selected from RMR-1029 using sample selection methods similar to the samples submitted to the FBI repository. | GAO-15-80

<sup>a</sup>Includes negative and inconclusive results as nonpositive results. The estimated negative rate is the number of non-positive results divided by the number of replications.

That’s a completely unacceptable result. The test called A1, when run 30 times in a row on material from the “smoking gun” RMR-1029, failed to detect the DNA variation in 13 of those tests. It gave a false negative in 43% of the tests when run on a known positive. And yet the FBI relied on this worthless test as part of the evidence to close the case.

### Exclusion of Samples With One Inconclusive Test

If reliance on a worthless test isn’t disturbing enough, the GAO report also dug out a point that was obscure in the NAS report. The FBI stated all along that in carrying out their analysis of the submitted cultures, they chose to eliminate from consideration any culture that gave an inconclusive result on any of the tests. But it turns out that there were some samples that definitely deserved further attention among those that were thrown out:

The NAS report also raised concerns that the decision to remove samples with inconclusive or variant results contributed to the lack of completeness of the repository data. The report stated that a major concern was the restriction of its statistical analyses to the 947 samples that contained no inconclusive or variant results. Notably, the report showed that 4 of the 112 samples that were disregarded for having a single inconclusive or variant result scored positive for the three remaining genetic tests.

Think about that for just a minute. Recall that only 8 of the 947 included samples tested positive for all four changes. And yet there are four more potential samples that might have all four DNA changes that have three positives and one inconclusive among the 112 that had an inconclusive result.

Going back to find that information in the NAS report makes it even worse. It turns out that among the 947 samples included in the final analysis, there were only three that had three positive tests, so the four with three positives and one inconclusive among the excluded 112 is huge. Here is a table with those four samples:

In addition to the two 3-positive samples (+++) among the 947 samples, the four samples below also tested positive for 3 mutations (ordered by FBIR number):

052-026	+	+	+	inc	-	A1, A3, MRI-D
053-010	var	+	+	+	+	A3, MRI-D, IITRI-D, E
054-008	inc	+	+	inc	+	A3, MRI-D, E
054-066	+	Inc	+	+	+	A1, MRI-D, IITRI-D, E

Where did samples 052-026, 053-010, 054-008 and 054-066 come from? The falsely closed Amerithrax investigation needs to be reopened to follow these sloppily discarded leads.