Clinical Chemistry 52:4
000–000 (2006)

Laboratory Management

Methodologic European External Quality Assurance for DNA Sequencing: The EQUALseq Program

Parviz Ahmad-Nejad,[1†] Alexandra Dorn-Beineke,[1†] Ulrike Pfeiffer,[1] Joachim Brade,[2] Wolf-Jochen Geilenkeuser,[3] Simon Ramsden,[4] Mario Pazzagli,[5] and Michael Neumaier[1*]

Background: **DNA sequencing is a key technique in molecular diagnostics, but to date no comprehensive methodologic external quality assessment (EQA) programs have been instituted. Between 2003 and 2005, the European Union has funded, as specific support action the EQUAL initiative to develop methodologic EQA schemes for genotyping (EQUALqual), quantitative PCR (EQUALquant), and sequencing (EQUALseq). Here we report on the results of the EQUALseq program.**

Methods: **The participating laboratories received a 4-sample set comprising 2 DNA plasmids, a PCR product, and a finished sequencing reaction to be analyzed. Data and information from detailed questionnaires were uploaded online and evaluated by use of a scoring system for technical skills and proficiency of data interpretation.**

Results: **Sixty laboratories from 21 European countries registered, and 43 participants (72%) returned data and samples. Capillary electrophoresis was the predominant platform (n = 39; 91%). The median contiguous correct sequence stretch was 527 nucleotides with considerable variation in quality of both primary data and data evaluation. The association between laboratory performance and the number of sequencing assays/year was statistically significant ($P < 0.05$). Interestingly, more than 30% of participants neither added comments to their data nor made efforts to identify the gene sequences or mutational positions.**

Conclusions: **Considerable variations exist even in a highly standardized methodology such as DNA sequencing. Methodologic EQAs are appropriate tools to uncover strengths and weaknesses in both technique and proficiency, and our results emphasize the need for mandatory EQAs. The results of EQUALseq should help improve the overall quality of molecular genetics findings obtained by DNA sequencing.**

In recent years, molecular biology–based methods have become widely used in clinical laboratories, and the latest genetic analyzers can perform a wide variety of sequencing and fragment analysis applications in healthcare, including microsatellite analysis, mutations detection, single-nucleotide polymorphism (SNP)[6] validation, screening, or even whole bacterial genome sequencing *(1–5)*.

DNA sequencing, regarded as the "gold standard" method, has high specificity and is routinely used in clinical diagnostics for mutational or SNP analysis. For example, direct sequencing analysis for the detection of mutations in the genes coding for cationic trypsinogen (*PRSS1*) or trypsin inhibitor (serine protease inhibitor, Kazal type I; *SPINK1*) in children with chronic pancreatitis *(6, 7)*, mutations in the cystic fibrosis gene *(8–10)*, or SNPs in the gene encoding coagulation factor VII *(11–13)* can have direct therapeutic consequences. In addition, the

[1] Institute for Clinical Chemistry and [2] the Department for Statistical Analysis, University Hospital Mannheim of the University of Heidelberg, D-68167 Mannheim, Germany.

[3] German Society for Clinical Chemistry and Laboratory Medicine (DGKL) Reference-Institute for Bioanalytics, Bonn, Germany.

[4] National Genetics Reference Laboratory (Manchester), St Mary's Hospital, Manchester, United Kingdom.

[5] Clinical Biochemistry Unit, Department of Clinical Physiopathology, University of Florence, 50139 Florence, Italy.

† These authors contributed equally to this work.

* Address correspondence to this author at: Institute for Clinical Chemistry, Theodor-Kutzer-Ufer 1-3, D-68167 Mannheim, Germany. Fax 49-621-383-3819; e-mail michael.neumaier@ikc.ma.uni-heidelberg.de.

Received October 5, 2005; accepted December 22, 2005.
Previously published online at DOI: 10.1373/clinchem.2005.061572

[6] Nonstandard abbreviations: SNP, single-nucleotide polymorphism; EQA, external quality assessment; and EU, European Union.

number of genetic factors investigated by sequencing techniques is growing steadily, and human genetics external quality assessment (EQA) schemes are being provided for some genetic defects in hereditary monogenic disorders e.g., by the European Molecular Genetics Quality Network (EMQN) *(14–18)*. In contrast, methodologic EQA programs for quality assessment have not been systematically established for DNA sequencing. Methodologic EQAs that address single aspects in preanalytical steps, DNA preparation and amplification, and reporting have been reported as early as 1998 *(19)*, and recommendations on quality assurance of molecular methods have been published by the IFCC *(20)*.

To investigate methodologic influences in the most common techniques in molecular diagnostics, the European Commission has funded, within the specific support action of life sciences, genomics and biotechnology for health, a project (EQUAL) to develop molecular EQAs for qualitative analysis, quantitative PCR, and DNA sequencing, designated EQUALqual, EQUALquant, and EQUALseq, respectively.

We developed EQUALseq focusing on different issues: (*a*) length and quality of DNA sequences obtained from plasmid DNA; (*b*) length and quality of DNA sequences obtained from a PCR product; (*c*) purification of a finished sequencing reaction and subsequent sequencing run; and (*d*) data interpretation. Here we present the results of this first pilot EQA study.

Materials and Methods

OVERALL DESIGN OF THE EQA

Participation was open to all laboratories carrying out sequencing-based methods within the European Union (EU) irrespective of their affiliation. The EQUALseq program was limited to 60 participating laboratories, and participation was free of charge. Registration was performed online via a dedicated website on the German Society for Clinical Chemistry and Laboratory Medicine (DGKL) homepage (http://www.dgkl-rfb.de/). Each participating laboratory was assigned an individual login password to be used for communication with the coordinator, for uploading results, and for retrieving both individual and general evaluation documents and information. Registration closed November 6, 2004, followed by shipment of the EQUALseq samples through Express carrier, which ended December 19, 2004. The survey closed on February 26, 2005.

With respect to the material shipped to the participants, we provided plasmid DNAs, a purified PCR product, and a finished sequencing reaction to be purified and analyzed (see section below).

Each laboratory received a set of 7 vials (3 containing primers, 4 with the samples), color-coded to avoid mistakes. Detailed instructions were enclosed in the vials and were also available online on the homepage of the European Communities Confederation of Clinical Chemistry (http://www.ec-4.org/equal/10_equal_seq.htm).

SAMPLE PREPARATION

Primers. Primers previously quality-controlled by mass spectrometric analysis were purchased from MWG-Biotech. Appropriate aliquots were dried and labeled.

Primers provided for sequencing were as follows: Sequencing primer 1 (100 pmol, dried pellet; 5′-AACAGCTATGACCATG-3′) to use in the sequencing reaction with sample 1; sequencing primer 2 (100 pmol, dried pellet; 5′-CATGGTCCTGCTGGAGTTCGTG-3′) to use in a sequencing reaction with sample 2; and sequencing primer 3 (100 pmol, dried pellet; 5′-GCCTCAGAGCAGGACCTTGG-3′) to use in a sequencing reaction with sample 3.

Plasmid DNAs. All plasmids used in this trial were transformed in DH5α and isolated from overnight cultures by use of the Nucleo-Bond PC 500 EF Kit (Macherey-Nagel). Quality was assessed by agarose electrophoresis, and quantity was assessed by spectrophotometry.

PCR product. Sample 3 was generated at the coordinator's laboratory by use of a DNA sample from a patient heterozygous for 2 coding mutations in the *HFE* gene (H63D and S65C) and for 1 homozygous mutation in intron 2 (IVS2 + 4) *(21, 22)*. This PCR product was generated in a reaction containing 2 μL of Primer FOR (10 pmol/μL; 5′-GCCTCAGAGCAGGACCTTGG-3′); 2 μL of Primer REV (10 pmol/μL; 5′-CAGCTGTTTCCTTCAAGATGC-3′); 5 μL of 10× buffer; 1.6 μL of 50 mM $MgCl_2$; 1 μL of 10 mM nucleotide mixture (Eppendorf); 0.4 μL of Platinum Taq DNA polymerase (5 U/μL; Invitrogen); 280 ng of patient DNA; and 34 μL of H_2O. The amplification was carried out on a Mastercycler gradient (Eppendorf) with initial denaturation at 95 °C for 120 s, followed by 35 cycles of 95 °C for 30 s, 62 °C for 30 s, and 72 °C for 30 s, and final elongation at 72 °C for 600 s. The amplicon was purified by use of a QIAquick PCR Purification Kit (QIAGEN) to remove the primers and byproducts. The concentration of the PCR product was determined by spectrophotometry, and sample 3 was aliquoted and subsequently dried as 300-ng aliquots.

SEQUENCING PERFORMED AT THE COORDINATOR'S FACILITY

Sequencing reactions were performed with the ABI PRISM® BigDye® Terminator v1.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems) as follows: 2 μL of Ready Reaction Mix; 1 μL of BigDye Seq. Buffer (5×); 0.25 μL of primer 1 (10 pmol/μL); 200 ng of plasmid DNA; and 4.75 μL of water. The sequencing reaction was run on a Mastercycler gradient (Eppendorf) with the following program: initial denaturation at 96 °C for 60 s, followed by 25 cycles of 96 °C for 10 s, 50 °C for 30 s, and 60 °C for 4 min. Sequencing reactions were subsequently purified by use of the DyeEx™ 2.0 Spin Kit (QIAGEN) according to the manufacturer's recommendations. A 4-μL portion of this purified probe was mixed with 16 μL of HiDi

buffer and run on an ABI PRISM 310 Genetic Analyzer equipped with a 47 cm × 50 μm (i.d.) 310 capillary filled with POP6 polymer (all from Applied Biosystems). The injection time was 30 s, and the running time was 50 min.

For sample 4, no purification of the sequencing reaction product was carried out because this sample was provided to the participants for purification and sequencing only.

SAMPLE FEATURES

Sample 1. The purpose of sample 1 was to check for the sequencing reaction performance of the participating laboratories, including purification and result reading. We provided 2 μg of a commercial cloning and sequencing vector (pCR TOPO2.1; Invitrogen) containing a 2355-bp insert. The sequencing primer was provided.

Sample 2. Sample 2 was a mixture of 2 different DNA plasmids (recombinant and nonrecombinant) generated to provoke comments or notes/interpretation from the participants with respect to template quality. Specifically, sequencing of sample 2 with primer 2 (GFP-specific sequence in the vector backbone pEGFP-C2; BD Clontech) produced ambiguous sequence reading downstream of the *Eco*R1 restriction site used for cloning of the 560-bp full-length cDNA of the murine small GTPase *Rab5*. As in sample 1, 2 μg of the plasmid mixture was provided to the participants together with sequencing primer 2.

Sample 3. Sample 3 was 300 ng of a purified 310-bp PCR product obtained from a patient with clinically significant heterozygous mutations in the *HFE* gene. A primer was sent for performing the reaction as well. Participants were asked to identify the sequence.

Sample 4. Sample 4 was generated from sample 1 as a finished sequencing reaction (see above). After purification and electrophoresis, the participants were asked to resend the remainder of the sample to the coordinator's laboratory for further investigation and reevaluation.

INSTRUCTIONS GIVEN

Participants received detailed instructions for actions to be performed on all materials to be analyzed. Concerning samples 1–3, the participants were asked to perform the sequencing reaction and to provide the longest possible sequence without any mistakes (e.g., "N"). It was explicitly stated not to perform primer-walking and to use exclusively the primer provided by the coordinator. Identification of the sequence and additional comments should be uploaded to the EQUALseq web page.

After analysis, sample 4 was to be sent back to the coordinator for further investigation. The BigDye Terminator v1.1 Cycle Sequencing Kit is compatible with sequencers such as the ABI PRISM 310 or ABI PRISM 3100. Participants using other sequence analyzers were not required to run sample 4 as part of the EQUALseq program. Again, identification and additional comments were requested.

SCORING OF PARTICIPANTS' RESULTS

Sample 1. To allow some degree of ranking, we developed a scoring system for technical and proficiency performance. Briefly, the sequencing reactions obtained at the coordinator's laboratory were taken as the 100% performance reference, starting at nucleotide 70 downstream of the primer annealing sites. The objective of sequencing the samples was to provide the longest possible sequence without any mistakes or ambiguities (i.e., N). Manual sequence editing and the use of database support were allowed.

A sequencing stretch of 525 nucleotides from the starting point (result obtained at the coordinator's laboratory) was taken as the 100% reference, on the basis of which the following basic and additional points were awarded to the participants:

- One basic point was awarded for every 50 bases of sequence read correctly. Scoring was terminated after the fifth sequencing error. The basic points were therefore awarded according to the total correct sequence length.
- Two additional points were awarded for sequences beyond the reference 525 bases without a single error or ambiguity.
- One additional point was awarded if the participant provided a shorter region, however, without any mistakes.
- One additional point was awarded for correct identification of the cloned insert in sample 1.

The final score was determined, and the number of mistakes (maximum of 5 points) was subtracted. For example, participant EQSRO068 provided a sequence of 718 nucleotides terminated after the fifth false base, representing 14 basic points. Because 524 bases were read without error, 2 additional points were granted. Because no identification was given, no additional point could be added. Five points were subtracted for the 5 false or ambiguous readings, giving a final score of 11 points for participant EQSRO068.

Sample 2. For sample 2, we awarded a maximum of 2 points. One point was given for a general comment on problems encountered with ambiguous readings, and an additional point was awarded for providing precise analysis of sample 2 based on the primary results.

Sample 3. We awarded a maximum of 6 points for correct analysis of sample 3: We gave 1 point for identification of the *HFE* PCR product, 1 point for detection of the structure of the PCR product (e.g., genomic DNA or cDNA), 1 point for a general comment on the deviation from the wild-type *HFE* sequence, 1 point for identification of the H63D mutation, 1 point for identification of the S65C

mutation, and 1 point for detection of the IVS2 + 4 variation.

Sample 4. Not every laboratory was able to analyze this sample because some sequencing platforms were incompatible with the chemistry used to generate sample 4. Evaluation was carried out as described for sample 1. Our own DNA sequencing results for sample 4 (no mistake for 480 bases beyond the *Eco*RI cloning site) were again taken as the 100% reference. Subsequent to purification and analysis, the purified sample 4 was requested for reanalysis in our laboratory. This allowed assessment of the purification efficiencies and relative performances of the sequencers compared with the single-capillary ABI PRISM 310 instrument used in our laboratory. Because of the low number of participants in this part of EQUALseq, no scores were calculated for sample 4.

EQUALseq questionnaire

A major objective was to compare instruments and technical procedures with the outcome in the sequencing program. We therefore asked the participants to complete a questionnaire before uploading their results to the EQUALseq web page (www.dgkl-rfb.de/). The data entry was divided into 3 sections: The first section was related to details of the methods used by the individual participants. The second part comprised questions about the equipment and procedures used. In addition, the participants were asked to give comments regarding the design of the EQUALseq program. The third part of the questionnaire was dedicated to data uploading.

All participants needed to register with the European Communities Confederation of Clinical Chemistry homepage (www.ec-4.org/equal/home.htm), thereby receiving a user identification for all further actions. All data were stored in a backend database (4^{Th} Dimension™). After closing of the data collection phase, the results were exported into a spreadsheet-compatible format. After evaluation of the individual data by Chromas (Ver. 1.45; Technelysium Pty Ltd), MacMolly® Tetra (Ver. 3.9; Soft Gene/Mologen Holding AG), and Microsoft Excel 2000, the results were presented on the DGKL website as a general report and as individual report sheets (PDF files). Individual reports were accessible only through the respective user identifications.

statistical analysis

Statistical differences between groups were tested by the Mann–Whitney *U*-test with a nonparametric approach (SAS™ software, Ver. 8.2; SAS Institute). P values <0.05 were considered significant.

Results

Of 76 requests to participate, 60 participants from 21 European countries were registered in EQUALseq. Of these, 43 laboratories (72%) from 18 countries (15 EU countries and 3 non-EU countries) completed the EQA and returned data and samples for evaluation to the coordinator. The affiliations included university hospital laboratories (n = 27), regional hospital laboratories (n = 9), research laboratories (n = 3), and private laboratories (n = 4). Seventeen laboratories dropped out of the EQA for various reasons. With respect to the analysis of sample 4, 12 participants (28%) did not participate in this aspect of the program because of incompatible sequencing platforms. Results of sample analysis are described in more detail below.

results for sample 1

The longest DNA sequence provided for sample 1 encompassed 894 nucleotides before termination after the fifth base error. Ten (23%) laboratories could determine 700 and more bases, using a single primer. The mean sequence stretch was 527 bases. Only 14 participants identified the insert correctly. However, we observed no association between the length of sequence and participant comments (see Fig. 1; also see Table 1 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol52/issue4/) that would have indicated that the best performances were in part attributable to working with genetic databases.

Also shown in Table 1 in the online Data Supplement are lists of all of the results; these results demonstrate the diversity of the data as obtained from the participants in both the length of sequence strings and the quality of managing of raw data. Five laboratories had a high frequency of errors or made no effort to edit ambiguous base calls leading to premature termination of sequence scoring. As extractable from the original ABI traces, editing of the raw data would have generated valuable sequence information in 3 of these cases, thereby giving a better score. This emphasizes the need to edit sequence data. Independent of the sequence length, only 14 (33%) laboratories provided clean sequences without any false bases (e.g., N). Eight (19%) participants reported ambiguous or false base calls within the first 250 bases. The sequence stretches and false bases are shown schematically in Fig. 1. It is apparent that most false bases/mistakes were generated at the end of the string, and no general difficulty could be determined on any sequencing platform with this sample.

results for sample 2

Sample 2 was included into EQUALseq to test the awareness of problems in raw data by providing a mixed, but partly identical template. We monitored comments concerning the nature and quality of sample 2 as well as explanations offered by the participants regarding the nature of the ambiguous sequence. Remarkably, 9 (21%) laboratories did not register any kind of difficulties in this part of the EQA, whereas the majority of participants commented on this sample. Indeed, excellent specific comments were received in some cases, not only correctly

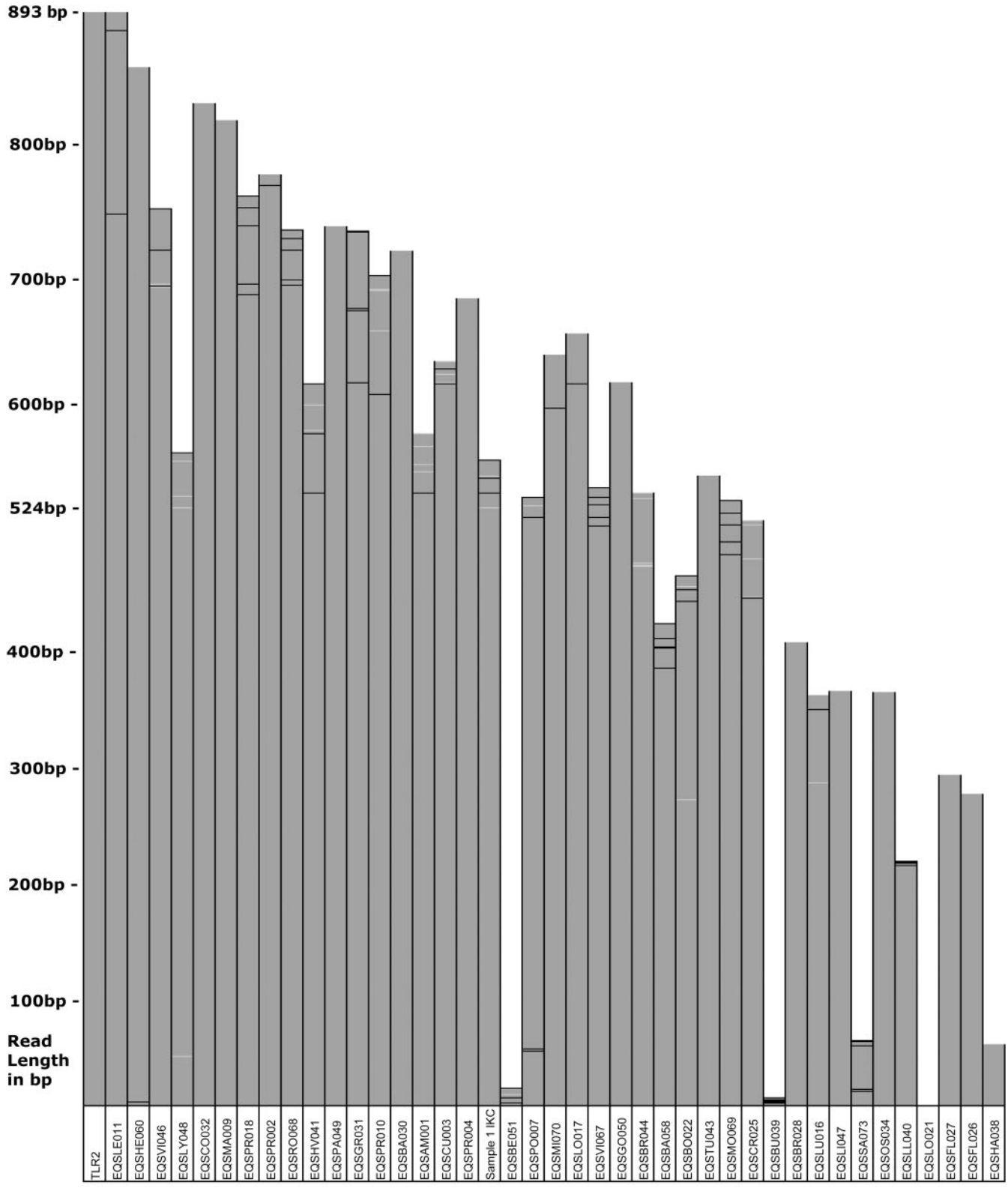

Fig. 1. Schematic presentation of the sequence strings for sample 1 as provided by the participants.

Dark gray represents accurate reading; *light gray* represents an "N" reading; *black* indicates the position of a n incorrect reading by comparison with Medline accession no. NM003264 *(24)*, which represents the gene for human toll-like receptor 2 (*TLR2*). *Sample 1 IKC* was defined as the reference basis for calculation of points. Sequence strings beyond the 5th false base are not displayed. Note the varying length and quality of the sequence strings and the distribution of errors.

interpreting the nature of the analytical problem but also identifying both plasmids by name. In one case of remarkable proficiency, the background sequence was subtracted, presumably by use of database information, to allow characterization of the cloned insert. In context with the data presented from sample 1, the results from sample 2 suggest that, although ~80% of laboratories will react to very obvious artifacts, single-base errors have a higher chance of going unnoticed.

RESULTS FOR SAMPLE 3

The PCR fragment representing sample 3 differed from the human wild-type *HFE* gene in 3 distinct nucleotide positions. It is noteworthy that 24 (57%) of the participants made no efforts for identification and that 27 (64%) and 26 (62%) of the participants did not comment on the heterozygous H63D or the S65C compound heterozygosity, respectively. However, the genotype coding for H63D and S65C was clearly present in 40% (n = 17) and 36% (n = 15) of the raw sequence datasets returned (Fig. 2). Accordingly, 13 (31%) of the 42 participating laboratories received no points in this part of EQUALseq, and only 4 (10%) laboratories were awarded all 6 points. One laboratory reported incorrect nucleotides at every position of the sample 3 sequence. This systematic mistake can be explained only by reporting of the reverse complement sequence in 3′ to 5′ orientation. Only then will every base be wrong. Taken together, we were surprised by the high proportion of reports of genetic variants without interpretation because the participants had been asked categorically for comments on the nature of sample 3 (see Table 2 in the online Data Supplement).

RESULTS FOR SAMPLE 4

The template of sample 4 was identical to that of sample 1. However, in contrast to sample 1, the participants were provided with a finished sequencing reaction (done in our laboratory), which subsequently had only to be purified and run on the sequencer. To address the influence of the participants' purification procedures, the purified samples were reanalyzed after being sent back from the laboratories.

Among 29 laboratories that were able to analyze both samples 1 and 4, 3 (10%) received the same score (see Table 1, whereas 15 (52%) participants performed better with sample 1, indicating that these laboratories delivered higher quality with their own sequencing procedures. In contrast, 11 (38%) laboratories gained more points with sample 4 than with sample 1. This may indicate that optimization of their enzymatic PCR sequencing reactions may be possible. Specifically, increases of more than 3 points (i.e., 150 bases) were possible in 13 (45%) laboratories. In 16 (55%) laboratories, the differences between both samples were less than 3 points. Another result of this part of EQUALseq was that transport of completed sequencing reactions at ambient temperature does not generally influence the quality of the DNA sequencing results (see Table 3 in the online Data Supplement).

The differences in ranking for the individual laboratories, as calculated with the EQUALseq scoring system, are shown in Figs. 3 and 4. Specifically, Fig. 3 refers to all laboratories, whereas Fig. 4 lists all participants that analyzed all 4 samples.

EQUALseq QUESTIONNAIRE

The first question in the EQUALseq questionnaire was the number of sequencing reactions performed per year by each laboratory. The largest group of participants (n = 14l 32.6%) performed 1000 to 4999 sequencing reactions per year (median, 2600). Overall, the minimum was 20, and the maximum was 40 000 sequencing reactions per year, with a median of 1600 (data not shown). The minimum template requirements differed from 1 to 1000 ng/μL of PCR product (median, 10 ng/μL) and 1 to 500 ng/μL of plasmid DNA (median, 100 ng/μL). The methods used (all nonradioactive dideoxy terminator methods) for sequencing were capillary sequencing by 39 participants (91%) and gel-based sequencing by 4 participants (9%).

To test whether the number of sequencing reactions performed per year possibly influenced the technical and analytical quality of the performance, we categorized the participants into 4 groups according to their number of sequencing assays per year and then assigned the points for samples 1 + 2 + 3 and sample 4 to the groups (Fig. 5). Participants running 1000–50 000 assays annually (n = 24) performed significantly better on samples 1 + 2 + 3 than participants running <1000 sequences per year (n = 19; $P = 0.04$). We also noted a trend for sample 4, but the differences did not reach statistical significance (n = 14 and n = 17; $P = 0.06$). These results suggest that experienced laboratories with higher throughput are better performers. The failure of these laboratories to reach statistical significance with sample 4 may be explained by the fact that they had to analyze a sample with an enzymatic sequencing step not optimized to their overall procedure.

Of the 43 participants, 37 (86%) are working with ABI PRISM Genetic Analyzers from Applied Biosystems Inc. Ten laboratories (23%) are working with a 1-capillary system, whereas 29 (67%) are working with multiple-capillary analyzers (Table 1). The median numbers (range) of annual sequencing assays performed on the different instruments are as follows: ABI PRISM 310 Genetic Analyzer, 500 (50–3000); ABI PRISM 377 Genetic Analyzer, 2700 (500–20 000); ABI PRISM 3100 Genetic Analyzer, 2750 (20–20 000); ABI PRISM 3700 Genetic Analyzer, 31 072 (120–40 000); ABI PRISM 3730 XL Genetic Analyzer, 30 000 (no range; used on only 1 participating laboratory); MegaBACE™ 1000, 12 516 (31–25 000); ALFexpress, 700 (no range; used on only 1 participating laboratory); CEQ™ 8000 Genetic Analysis System, 3500 (200–4200). Users working with high-throughput analyzers (n = 29) performed a mean of 3500 sequencing reactions per year

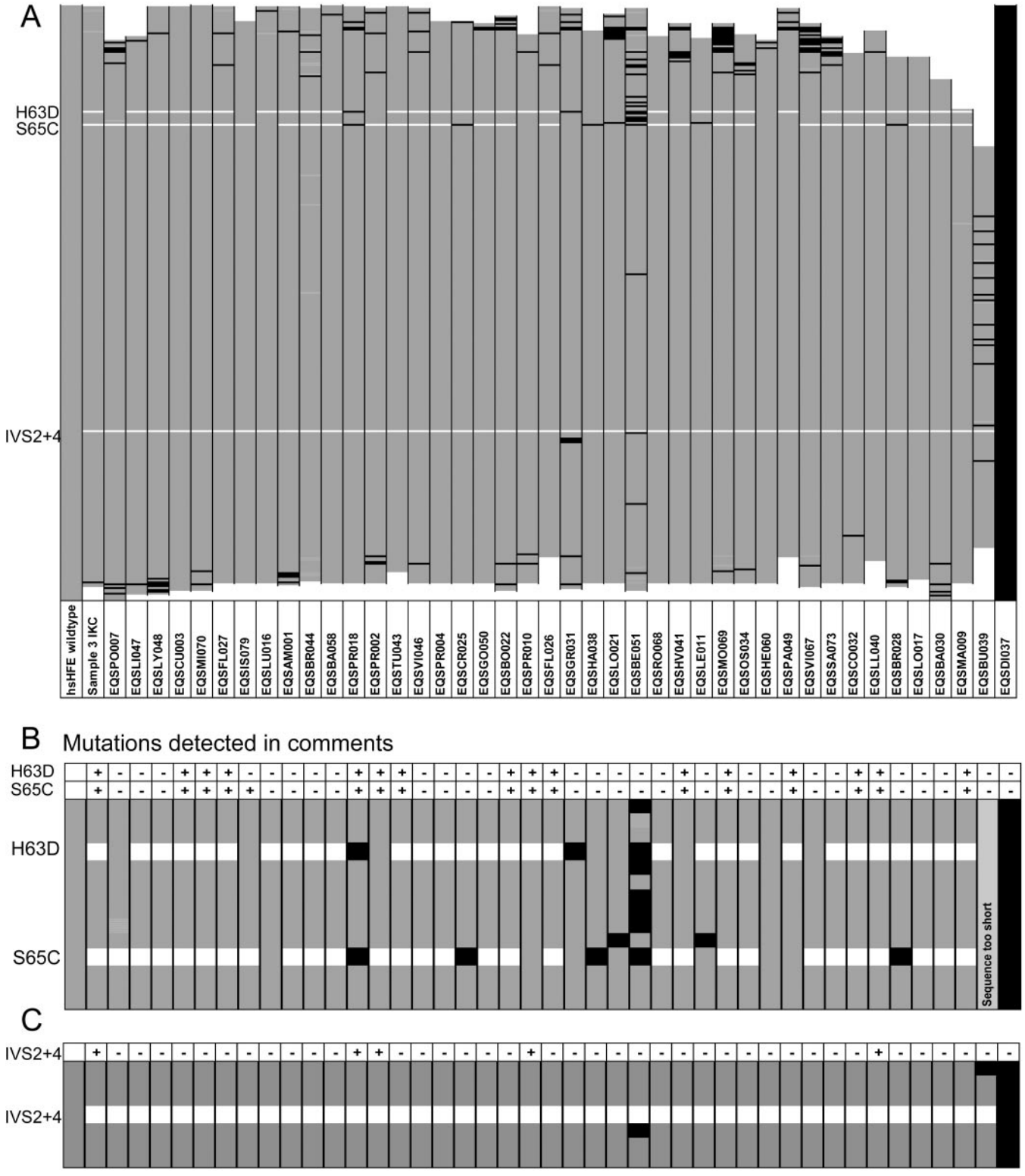

Fig. 2. Schematic representation of the DNA sequence data provided for the 310-bp PCR fragment (Sample 3).

(*A*), schematic of the overall sequence string obtained from each participant; (*B* and *C*), regions of interest. Color codes are as follows: ▨, nucleotide position is identical to Medline accession no. Z92910, wild-type human *HFE* gene; □, correct reading of the sequence at the positions of known SNPs in *HFE*; ■, sequence error reported by participant or mutation not detected; ▤, ambiguous "N" reading reported by participant. + and − indicate the presence or absence of comments on detection of the H63D, S65C, or IVS2 + 4 SNPs present in sample 3.

Table 1. Instruments used by the laboratories participating in the EQUALseq program and related scores for samples 1 + 2 + 3 and sample 4.[a]

	Total score for sample 1 + 2 + 3				Score for sample 4			
DNA sequencing instrumentation used	n	%	Median	Range[b]	n	%	Median	Range
ABI PRISM 310 Genetic Analyzer[c,f]	10	23	10.0	−5.00 to 21.0	8	19	5.50	3.00–15.0
ABI PRISM 377 Genetic Analyzer[c,f]	3	7	6.00	−3.00 to 8.00	3	7	4.00	0.00–14.0
ABI PRISM 3100 Genetic Analyzer[c,g]	20	47	14.5	0.00–25.0	16	37	8.50	−5.00 to 17.0
ABI PRISM 3700 Genetic Analyzer[c,g]	3	7	24.0	21.0–26.0	2	5	17.5	17.0–18.0
ABI PRISM 3730 XL Genetic Analyzer[c,g]	1	2	0.00		1	2	−4	
MegaBACE1000[d,g]	2	5	19.5	18.0–21.0	1	2	16.0	
ALFexpress[d,f]	1	2	12.0		0	0		
CEQ8000 Genetic Analysis System[e,g]	3	7	5.00	3.00–15.00	0	0		
Overall	43	100	12.0	−5.00 to 26.0	31	72	8.00	−5.00 to 18.0

[a] Participants are categorized in groups according to the instruments used.
[b] Minimum–maximum.
[c] Applied Biosystems (Darmstadt, Germany).
[d] Amersham Pharmacia Biotech, GE Healthcare (Freiburg, Germany).
[e] Beckman Coulter (Krefeld, Germany).
[f] Low-throughput analyzer.
[g] High-throughput (multiple-capillary) analyzer.

(range, 20–40 000), whereas low-throughput users (n = 14) performed 600 (50–20 000) assays per year (P = 0.0178). The results for the instruments used and the scores for samples 1 + 2 + 3 and sample 4 are shown in Table 1. The overall median (ranges) numbers of points awarded were 12 (−5 to 26) for samples 1 + 2 + 3 (n = 43) and 8 (−5 to 18) for sample 4 (n = 31). Because of the small sizes of most of the groups, the participants were divided into 2 groups: those who used low-throughput analyzers, and those who used multiple-capillary analyz-

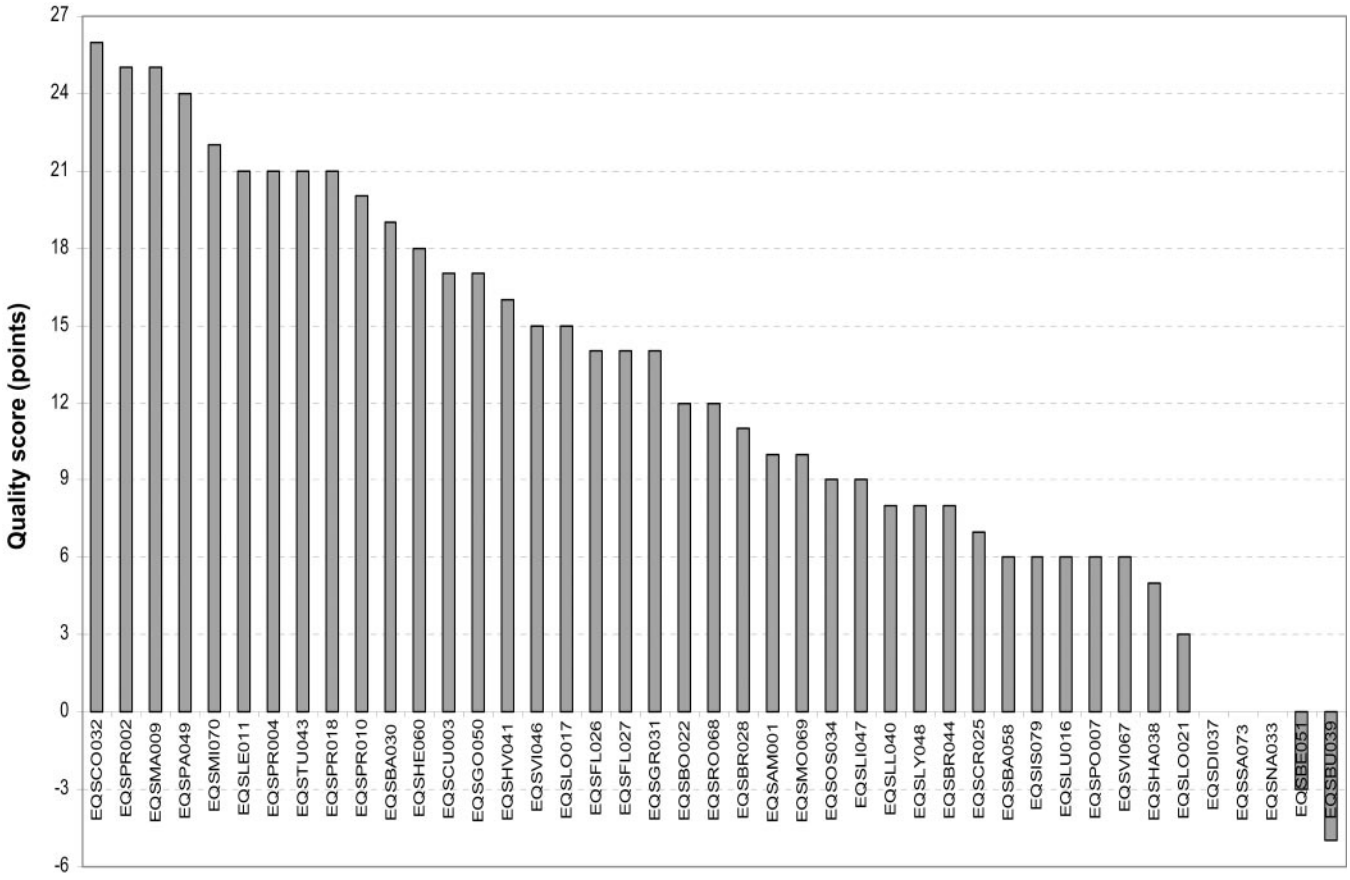

Fig. 3. Ranking of all EQUALseq participants that returned data (n = 43) for DNA samples 1–3.

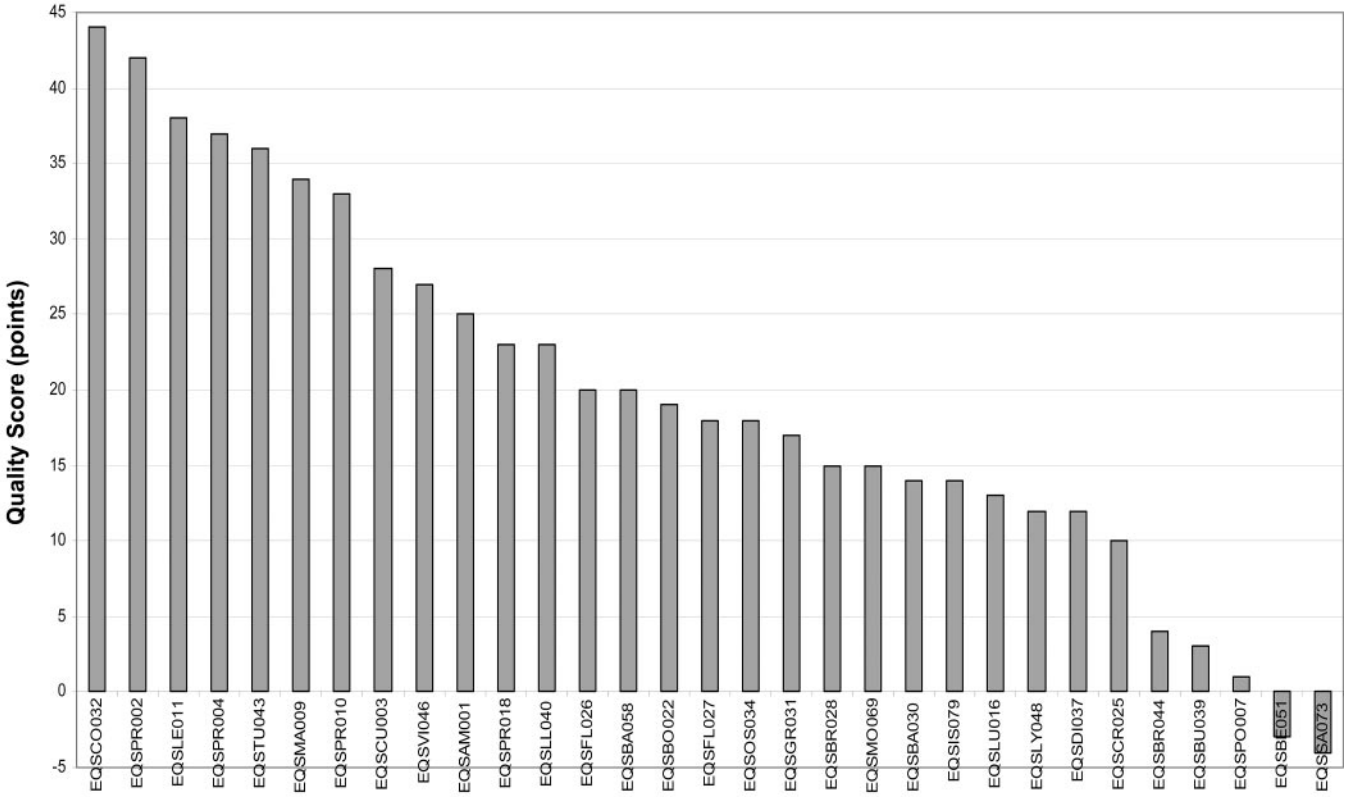

Fig. 4. Ranking of EQUALseq participants (n = 31) participating in the analysis of all 4 DNA samples.

ers. Regarding the points obtained from samples 1 + 2 + 3 and from sample 4, the medians (ranges) were as follows: 9 (−5 to 21) and 5 (0–15.9) for users of low-throughput analyzers and 15 (0–26) and 9 (−5 to 18) for users of multiple-capillary analyzers. The differences in points awarded between the groups were significant for samples 1 + 2 + 3 (n = 14 and n = 29, respectively; $P = 0.01$) but not significant for sample 4 (n = 11 and n = 20, respectively; $P = 0.25$). The results indicate that the number of sequencing assays and the quality of the enzymatic PCR sequencing established within each participant's laboratory are important variables for predicting good sequencing performance. The national origins of the laboratories did not have an impact on quality, whereas the affiliations of the laboratories did: regional hospital laboratories seemed to perform somewhat more weakly than the university or research laboratories. In addition, the few private laboratories entered in this study performed well (see Fig 1 in the online Data Supplement).

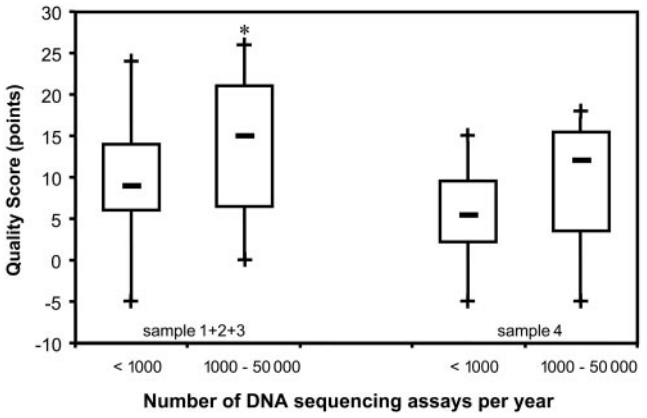

Fig. 5. Overall quality of DNA sequencing as a function of the number of sequencing assays performed per year.

Box and whisker plots depicting the medians (lines inside boxes) and minima and maxima (*error bars*) of the scores. The participants were categorized into 2 groups depending on their annual sequencing throughput. The results are displayed separately for samples 1–3 (n = 43) and sample 4 (n = 31). Participants performing 1 000–50 000 sequencing assays per year (n = 24) performed significantly better in sample 1–3 than participants performing <1 000 sequencing assays per year (n = 19; $P = 0.04$). Performance of those laboratories that analyzed sample 4 was not statistically different between the 2 groups (n = 14 and n = 17; $P = 0.06$).

Table 2 gives an overview of the reagents sets used for sequencing and lists the results from the descriptive statistical analysis. Because of the small sample size of some groups, nonparametric testing with the Kruskal–Wallis test was not appropriate. We therefore tested the groups with sample sizes ≥3, using the Mann–Whitney *U*-test group by group. No statistically significant results were obtained for samples 1 + 2 + 3. Trend for better results were seen only for laboratories using the BigDye Terminator v3.1 Ready Reaction Cycle Sequencing Kit (Applied Biosystems) and the DYEnamic ET Dye Terminator Cycle Sequencing Kit (Amersham Pharmacia Bio-

tech, GE Healthcare). According to the manufacturer's statement, the BigDye Terminator v3.1 Ready Reaction Cycle Sequencing Kit is more suitable for longer sequence readings than is Ver. 1.1.

We also evaluated the methods for the removal of unincorporated dye terminators. A majority of participants (n = 17; 40%) preferred the ethanol precipitation method with generally good results, whereas the remainder used commercially available methods (Table 3). Again, because of the small sample size of some groups, nonparametric testing using the Kruskal–Wallis test was not appropriate, and testing was carried out with the Mann–Whitney *U*-test group by group as described above. The following significant results were obtained for samples 1 + 2 + 3: Sephadex vs Millipore (both n = 6; $P = 0.04$), Sephadex vs QIAGEN DyeEx (both n = 6; $P = 0.01$), and Sephadex (n = 6) vs Centri SEP™ Spin columns (n = 3; $P = 0.04$). No statistically significant differences were obtained for sample 4. The removal of unincorporated dye terminators with Sephadex seems to provide better results.

Finally, the participants were asked, after the retrieval of their results from the web page, to give some comments and suggestions about the EQUALseq program on completion of the program. Thirty-six participants (84%) gave one or more comments, and 21 participants (49%) made some specific recommendations. Twenty-six participants (72%) regarded the EQUALseq program as "interesting" or "very interesting". Special comments included statements such as "interesting is the comparison with other labs" (8%), "it was the first quality control concerning DNA sequencing in our lab" (22%), "it was interesting to use plasmid-DNA" (6%), and "more than ABI PRISM systems should be tested" (6%), a statement which referred to sample 4. The performance of participants according to their affiliations is shown online in Fig 1 of the online Data Supplement. No multivariate analysis was done because of the small sample size of some groups. The descriptive statistics indicate that private laboratories seem to perform better than university, research, and regional hospital laboratories.

Discussion

DNA sequencing is a key technique to definitely determine genetic sequences, and it is regarded as the gold standard of molecular analysis. In the postgenomic era, increasing numbers of polymorphisms and mutations are being detected and linked to specific diseases. Recent improvements to the technical side include the introduction of multiple capillary sequencing with up to 96 capillaries run simultaneously. This clearly demonstrates that data output is increasing at an enormous rate. New technologies that combine technologies such as digital PCR with Pyrosequencing are bring introduced, and results for "whole-genome sequence sets" of 20 million bases could be available in a matter of hours *(4)*. Although DNA sequencing is the traditional method of choice in human genetics applications investigating monogenic disorders, short-range sequencing may also become a method of choice for laboratory medicine, e.g., in the detection of SNPs associated with multifactorial diseases.

DNA sequencing is being routinely used in diagnostic EQA schemes for human genetics, e.g., through quality-control networks such as the European Molecular Genetics Quality Network (EMQN) *(14, 15)*. However, except for single reports *(23)*, dedicated methodologic EQA programs for DNA sequencing that concentrate on various aspects of analytical proficiency have not been initiated to the best of our knowledge.

The European EQUALseq program was designed to address methodologic procedures and analytical proficiency; 60 laboratories registered to the trial. Next to the generation of extended DNA sequences, the program encouraged the editing of raw data and interpretation as a matter of quality in the overall process. In addition, to evaluate both qualities within a DNA sequencing assay, we tried to devise a model scoring system to rank the participating laboratories in this methodologic DNA sequencing EQA. This scoring system judges the generation of long sequence strings together with the quality of manual editing of raw data and, to a lesser extent, identification of the samples. Medical interpretation or medical proficiency testing was not the scope of this methodologic EQA.

As one of the most striking results of EQUALseq we noticed a considerable diversity in sequencing performance, although 39 of 43 laboratories used capillary technology, of which 87% (n = 34) were using Applied Biosystems instrumentation and chemistry (Fig. 1 and Tables 2 and 3). In part, these may be explained by the extraordinary difference in the amounts of template that the laboratories used for setting up their enzymatic reactions. Specifically, we found that the participants routinely used between 1 and 1000 ng/μL (median, 10 ng/μL) PCR product or between 1 and 500 ng/μL (median, 100 ng/μL) plasmid DNA to generate PCR products for sequencing runs. This suggests different degrees of optimization between laboratories using the same platform. Similarly, we found that different methods were used for cleaning up the reactions. Specifically, there were differences between classic ethanol or 2-propanol precipitation methods and commercial reagents for purification of sequencing reactions (e.g., QIAGEN or Millipore; see Table 3), possibly leading to loss of signal intensity in some cases.

DNA sequencing is a complex multistep procedure being performed with different degrees of optimization in individual laboratories. To assess this issue, we provided the participants with both a DNA template to be run through the entire process of sequencing, purification, and electrophoresis (sample 1) and a prefabricated sequencing reaction (previously generated from sample 1 in our laboratory) to be analyzed on their instruments (sample 4). In this experiment. 14 laboratories performed better

Table 2. Reagents used by the laboratories participating in the EQUALseq program and related scores for samples 1 + 2 + 3.[a]

	Participants		Total score for samples 1 + 2 + 3	
Reagents used for DNA sequencing	n	%	Median	Range
BigDye Terminator v1.1 Ready Reaction Cycle Sequencing Kit[c]	10	23	10.0	6.00–26.0
BigDye Terminator v3.1 Ready Reaction Cycle Sequencing Kit[c]	16	37	14.5	−5.00 to 25.0
BigDye Terminator v1.1 and v3.1 Ready Reaction Cycle Sequencing Kit[c]	3	7	17.0	10.0–24.0
BigDye Terminator reagents, not specified[c]	5	12	8.00	0.00–14.0
dRhodamine Terminator Cycle Sequencing Ready Reaction Kit[c]	1	2	10.0	
DYEnamic ET Dye Terminator Cycle Sequencing Kit[d]	3	7	21.0	6.00–22.0
TempliPhi™ DNA Sequencing Template Amplification Kit and DYEnamic ET Dye Terminator Cycle Sequencing Kit[d]	1	2	18.0	
CEQ DTCS Quick Start Kit[e]	4	9	8.50	3.00–15.0

[a] Participants are categorized into groups according to the reagents used for sequencing.
[b] Minimum–maximum.
[c] Applied Biosystems.
[d] Amersham Pharmacia Biotech, GE Healthcare.
[e] Beckman Coulter.

with their own procedure (i.e., sample 1 gave superior results than sample 4), whereas 12 laboratories performed better with sample 4 than with sample 1. The results obtained from samples 1 and 4 allowed us to dissect and compare the performances of single steps within the sequencing procedure within this EQA. We suggest that such information may provide help to laboratories to optimize their sequencing efforts and therefore should be included into future EQAs (see Table 3 in the online Data Supplement). Finally, routine experience appeared to be important for the laboratories in the overall performance in EQUALseq: participants with more than 1000 sequencing assays per year were significantly better performers.

In contrast to judging the technical aspects of raw data, we consider it difficult to judge analytical proficiency in DNA sequencing. However, we were surprised to find that many laboratories restricted their attention to the purely technical issues of the electrophoretic run. Although asked to do so in EQUALseq, editing of the raw sequence data appeared not to be understood as a part of the technical expert's assessment among many participants. This explains why many laboratories that showed good technical performance fell short in their overall scores. We believe that improving what we tentatively would call "postanalytical proficiency" should be a major focus of future EQAs in DNA sequencing. Certainly, the shear amount of data generated by the present technology will require increased efforts, e.g., double-checking results and consulting of databases. The scoring system we have developed in the EQUALseq program was intended to be a first step toward quality assessment of the technical validation process. As can be seen in Figs. 3 and 4, the differences among the 43 laboratories entered into this evaluation were considerable. National origin of the laboratories did not have an impact on quality, whereas the affiliations of the laboratories did: regional hospital labo-

Table 3. Comparison of methods for purification of sequencing reaction products and related scores for samples 1 + 2 + 3 and sample 4.[a]

	Total score for samples 1 + 2 + 3				Score for sample 4			
Method for removal of unincorporated dye terminators from sequencing reaction	n	%	Median	Range[b]	n	%	Median	Range
Ethanol	17	40	14.0	0.00–26.00	12	28	10.5	−5.00 to 18.0
Sephadex	6	14	20.0	6.00–24.0	3	7	14.0	−5.00 to 16.0
Millipore	6	14	8.50	0.00–21.0	5	12	9.00	−4.00 to 17.0
Isopropanol	2	5	10.5	10.0–11.0	2	5	9.50	4.00–15.0
QIAGEN DyeEx[c]	6	14	8.00	−3.00 to 14.0	6	14	5.50	0.00–8.00
Centri SEP Spin columns[d]	3	7	7.00	−5.00 to 17.0	3	7	8.00	3.00–11.0
CleanSEQ™ dye terminator removal[e]	1	2	15.0		0	0		
Others	2	5	6.00	0.00–12.0	0	0		

[a] Participants are categorized into groups according to the methods used for removal of unincorporated dye terminators.
[b] Minimum–maximum.
[c] QIAGEN GmbH (Hilden, Germany).
[d] Princeton Separations (Adelphia, NJ).
[e] Agencourt (Beverly, MA).

ratories seemed to perform somewhat more weakly than the university or research laboratories. In addition, the few private laboratories entered in this study performed well (see Fig. 1 in the online Data Supplement). We want to emphasize that this diversity is not attributable to differences in the methods used because the technology used to complete EQUALseq was very standardized.

In conclusion, the European EQUALseq program shows that the methodologic aspects of DNA sequencing are crucial to uncover strengths and weaknesses in techniques and postanalytical proficiency. In our opinion, this program has demonstrated that there is considerable room for improvement. With respect to the increasing importance of DNA sequencing for clinical laboratory diagnostics, we propose that methodologic EQAs should be made mandatory for laboratories providing this service. The experiences obtained in EQUALseq may serve as an approximation to such future EQAs.

The EQUALseq program is part of the EQUAL program, which is funded by the 6th framework program of the European Commission.

References

1. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 1977;74:5463–7.
2. Langaee T, Ronaghi M. Genetic variation analyses by pyrosequencing. Mutat Res 2005;573:96–102.
3. Franca LT, Carrilho E, Kist TB. A review of DNA sequencing techniques. Q Rev Biophys 2002;35:169–200.
4. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005;437:376–80.
5. Marziali A, Akeson M. New DNA sequencing methods. Annu Rev Biomed Eng 2001;3:195–223.
6. Drenth JP, te Morsche R, Jansen JB. Mutations in serine protease inhibitor Kazal type 1 are strongly associated with chronic pancreatitis. Gut 2002;50:687–92.
7. Howes N, Greenhalf W, Rutherford S, O'Donnell M, Mountford R, Ellis I, et al. A new polymorphism for the RI22H mutation in hereditary pancreatitis. Gut 2001;48:247–50.
8. Ravnik-Glavac M, Glavac D, di Sant' Agnese P, Chernick M, Dean M. Cystic fibrosis gene mutations detected in hereditary pancreatitis. Pflugers Arch 1996;431:R191–2.
9. Witt H. Gene mutations in children with chronic pancreatitis. Pancreatology 2001;1:432–8.
10. Cuppens H, Cassiman JJ. Solid phase fluorescent sequencing of the CFTR gene. Methods Mol Biol 2001;167:63–88.
11. Giansily-Blaizot M, Aguilar-Martinez P, Biron-Andreani C, Jeanjean P, Igual H, Schved JF. Analysis of the genotypes and phenotypes of 37 unrelated patients with inherited factor VII deficiency. Eur J Hum Genet 2001;9:105–12.
12. Miller CH, De Staercke C, Benson J, Hooper WC, Dilley A, Evatt BL, et al. Elevated factor VII as a risk factor for recurrent fetal loss: relationship to factor VII gene polymorphisms. Thromb Haemost 2005;93:1089–94.
13. Ariffin H, Millar DS, Cooper DN, Chow T, Lin HP. Prenatal exclusion of severe factor VII deficiency. J Pediatr Hematol Oncol 2003;25:418–20.
14. Losekoot M, Bakker B, Laccone F, Stenhouse S, Elles R. A European pilot quality assessment scheme for molecular diagnosis of Huntington's disease. Eur J Hum Genet 1999;7:217–22.
15. Mueller CR, Kristoffersson U, Stoppa-Lyonnet D. External quality assessment for mutation detection in the BRCA1 and BRCA2 genes: EMQN's experience of 3 years. Ann Oncol 2004;15(Suppl 1):I14–7.
16. Muller CR. Quality control in mutation analysis: the European Molecular Genetics Quality Network (EMQN). Eur J Pediatr 2001;160:464–7.
17. Simoni M, Bakker E, Krausz C. EAA/EMQN best practice guidelines for molecular diagnosis of y-chromosomal microdeletions: state of the art 2004. Int J Androl 2004;27:240–9.
18. Ibarreta D, Elles R, Cassiman JJ, Rodriguez-Cerezo E, Dequeker E. Towards quality assurance and harmonization of genetic testing services in the European Union. Nat Biotechnol 2004;22:1230–5.
19. Braun A, Deufel T, Geilenkeuser WJ, Neumaier M, Rohle G, Roscher A, et al. External quality assessment of molecular biology-based methods used in laboratories of clinical chemistry and human genetics. Clin Chem Lab Med 1998;36:231–4.
20. Neumaier M, Braun A, Wagener C. Fundamentals of quality assessment of molecular amplification methods in clinical diagnostics. International Federation of Clinical Chemistry Scientific Division Committee on Molecular Biology Techniques. Clin Chem 1998;44:12–26.
21. de Diego C, Murga MJ, Martinez-Castro P. Frequency of HFE H63D, S65C, and C282Y mutations in patients with iron overload and controls from Toledo, Spain. Genet Test 2004;8:263–7.
22. Floreani A, Navaglia F, Basso D, Zambon CF, Basso G, Germano G, et al. Intron 2 [IVS2, T-C +4] HFE gene mutation associated with S65C causes alternative RNA splicing and is responsible for iron overload. Hepatol Res 2005;[Epub ahead of print].
23. Neumaier M, Braun A, Gessner R, Funke H. Experiences with external quality assessment (EQA) in molecular diagnostics in clinical laboratories in Germany. Working Group of the German Societies for Clinical Chemistry (DGKC) and Laboratory Medicine (DGLM). Clin Chem Lab Med 2000;38:161–3.
24. Rock FL, Hardiman G, Timans JC, Kastelein RA, Bazan JF. A family of human receptors structurally related to Drosophila Toll. Proc Natl Acad Sci U S A 1998;95:588–93.