Transfusion Medicine and Hemotherapy

Research Article

Transfus Med Hemother 2022;49:240–249
DOI: 10.1159/000521604

Received: September 27, 2021
Accepted: December 20, 2021
Published online: January 31, 2022

Non-Invasive Fetal K Status Prediction: 7 Years of Experience

Klaus Rieneck[a] Frederik Banch Clausen[a] Thomas Bergholt[b]
Lone Nikoline Nørgaard[b] Morten Hanefeld Dziegiel[a, c]

[a]Department of Clinical Immunology, Rigshospitalet, Copenhagen, Denmark; [b]Department of Obstetrics, Rigshospitalet, Copenhagen, Denmark; [c]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

Keywords

Kell blood group · Prenatal diagnosis · Hemolytic disease of the fetus and newborn · Non-invasive prediction · Cell-free DNA · Genotyping · Next-generation sequencing

Abstract

Introduction: In the Kell blood group system, the K and k antigens are the clinically most important ones. Maternal anti-K IgG antibodies can lead to the demise of a K-positive fetus in early pregnancy. Intervention can save the fetus. Prenatal K status prediction of the fetus in early pregnancy is desirable and gives a good basis for pregnancy risk management. We present the results from 7 years of clinical experience in predicting fetal K status as well as some theoretical considerations relevant for design of the assay and evaluation of results. ***Methods:*** Blood was collected from 43 women, all immunized against K, at a mean gestational age of 18 weeks (range 10–38). A total of 56 consecutive samples were tested. The *KEL*01.01/KEL*02* single nucleotide variant that determines K status was amplified from maternal plasma DNA by PCR without allele specificity. The PCR product was sequenced by NGS technology, and the number of sequenced *KEL*01.01* and *KEL*02* reads were counted. Prediction of the fetal K status was based on this count and was compared with the serologically determined K status of the newborns. ***Results:*** All fetal K predictions were in accordance with postnatal serology where available (n = 34), using our current data analysis. ***Conclusion:*** We have developed an NGS-based method for the non-invasive prediction of fetal K status. This approach requires special considerations in terms of primer design, stringent preanalytical sample handling, and careful analytical procedures. We analyzed samples starting at GA 10 weeks and demonstrated the correct prediction of fetal K status. This assay enables timely clinical intervention in pregnancies at risk of hemolytic disease of the fetus and newborn caused by maternal anti-K IgG antibodies.

Introduction

Alloimmunization against the K blood group antigen (ISBT KEL1 006001) is an important cause of hemolytic disease of the fetus and newborn with life-threatening fetal anemia [1, 2]. There is presently no possibility of specific Kell immune prophylaxis. Anti-K antibodies may persist in the newborn long after birth and necessitate blood transfusions.

Prediction of fetal K status in early pregnancy in women with anti-K alloantibodies enables focusing on the relatively few K-positive fetuses, with close monitoring with ultrasound and treatment. If the fetus is negative, the number of clinical and serological controls during pregnancy can be reduced.

In cases of anti-K immunized women carrying a K-positive fetus, the fetus may be severely affected by anemia or die in the second trimester without intervention [2]. Therefore, it is important to predict fetal K-status early in pregnancy to consider early treatment, especially in cases with prior severely affected pregnancies. Treatment options include weekly intravenous immunoglobins (IVIG) to pre-

karger@karger.com
www.karger.com/tmh

Correspondence to:
Klaus Rieneck, klaus.rieneck@regionh.dk

vent fetal hemolysis [3]. Ongoing studies are testing new forms of maternal immunomodulation (https://clinicaltrials.gov/ct2/show/NCT03842189). Standard treatment once fetal anemia is suspected is intrauterine transfusion (IUT) of the fetus with K-negative blood that should also be negative for other commonly immunogenic blood group antigens the woman does not have herself [4].

The K phenotype has its genetic basis in a single nucleotide variant (SNV) [5]. Prenatal K prediction from cell-free DNA (cfDNA) from maternal plasma is based on the detection of a single base substitution (rs8176058) in a small number of fetal cfDNA molecules in a large surplus of maternal cfDNA that encodes the antithetical antigen. To diagnose *KEL*01.01*-positive fetuses non-invasively early in the pregnancy, specific DNA amplification of fetally derived cfDNA in maternal plasma has been used by others [6]. One real-time PCR approach requires modified PCR primers to achieve a high degree of specificity for the *KEL*01.01* allele necessary to detect the fetal allele [7]. This is technically demanding, and early diagnosis may be difficult to achieve. Recently, others have used digital PCR that appears very promising [8], allowing precise estimation of fetal fraction, absolute measurement of K+ k+ versus k+ k+ copies/μL and circumvents some problems of conventional PCR. Others have shown that in a next-generation sequencing (NGS)-based approach, all fetomaternal incompatibilities were correctly detected in a small cohort [9].

We have implemented NGS technology in a robust approach to confirm or exclude the presence of *KEL*01.01* in the plasma of pregnant K-negative women with anti-K IgG antibodies from early pregnancy starting at gestational age (GA) 10 weeks [3, 10, 11]. The assay would also be useful to predict a k-positive fetus in a k-negative woman. When designing PCR primers for amplification of fetal cfDNA carrying genetic variation, keeping the gene-specific amplicon short remains a priority. The average length of fetal and maternal cfDNA is 143 and 166 bp, respectively [12]. The minimal length of a cfDNA fragment to be amplified must accommodate the binding of both specific primers.

The aim of this work was to make a reliable assay for clinical use for early antenatal fetal K status prediction. We present some experimental and some important theoretical considerations for this NGS-based approach as well as the consecutive data from 7 years of experience of predicting the fetal K status, including 6 samples from GA 10–12 weeks in K-immunized pregnant women.

Materials and Methods

Blood Samples

A total of 56 samples were collected from anti-K-immunized pregnant women and analyzed in the period 2013 to 2020 at the Blood Bank, Rigshospitalet, Copenhagen, Denmark. Age, anti-K titer, and GA are described in Table 1. Eleven samples were repeat samples from the same pregnancy; one sample was repeated three times. Blood was drawn into Cell-Free DNA BCT tubes (Streck Inc., La Vista, NE, USA) and stored at room temperature. The blood sample was centrifuged at 2,000 *g* for 10 min at room temperature as soon as possible after venipuncture. The plasma was separated from the pellet within 2 h of the blood draw in most cases, but blood samples from other hospitals were up to a few days old and shipped at room temperature. Plasma was carefully aspirated avoiding all leukocytes and placed in a new tube and kept at −20°C if not processed the same day.

The samples were taken at different GAs, six of which were from week 10–12. In total 6 women had IUT, and of these 4 had an anti-K titer of 1,024 or higher, with a predicted K-positive fetus. Two women had IUT for other reasons than anti-K: one had a high anti-D titer and no specific cause for the fetal anemia was found in the other case.

DNA Purification

cfDNA was purified from 4 mL of plasma using a QIAsymphony robot (Qiagen, Hilden, Germany), using a DSP Virus/Pathogen Midi-kit (ref. No. 937,055) as previously described [13]. The purified cfDNA was eluted into 60 μL of AVE buffer.

PCR Primers

Due to the small size of cfDNA, it is an important constraint to keep the amplicon short. We estimated the fraction of amplifiable cfDNA molecules as a function of amplicon size from theoretical calculations and used published experimental data (online suppl. Fig. 1a, b; see www.karger.com/doi/10.1159/000521604 for all online suppl. material) [14]. Assuming a random fragmentation of fetal genomic DNA into cfDNA fragments with an average size of 143 bp, the percentage of cfDNA allowing for PCR amplification can be estimated. The distribution of the experimental cfDNA data was visualized (online suppl. Fig. 1b). The gene-specific primers were synthesized with appended adaptors for an easier procedure. The sequences of the PCR primers are shown in online supplementary Table 1. The gene-specific amplicon size was 58 bp.

By amplifying and sequencing two bases in the *KEL* gene, the primers interrogate rs8176058 *KEL*01.01* (c.578C>T, p.Thr193Met) and the antithetical allele *KEL*02* (c.578C, p.Thr193), as well as rs61729031 *KEL*01.02* (c.577T>A, p.Thr193Ser), which can result in a weak K-positive phenotype. Also, we used the frequency of genetic variation from the gnome database (https://gnomad.broadinstitute.org/, June 10, 2020) to investigate the potential risk of allelic dropout caused by SNVs in the primer recognition sequence.

PCR Amplification

We used 16 μL of eluted cfDNA as the template for the PCR amplification. Thus, the overall plasma equivalent per PCR was 1.07 mL. The PCR was performed on an MJ Thermocycler (Bio-Rad, Hercules, CA, USA) or a Veriti Thermocycler (Applied Biosystems, Waltham, MA, USA) in a PCR with a total volume of 40 μL with 0.5 μM of each primer (Eurofins Genomics, Ebersberg, Germany), 1× Phusion master mix (New England Biolabs, Ipswich, MA, USA), and with the following thermoprofile: one cycle of 98°C 30 s, 40 cycles of 98°C 15 s, 63°C 15 s, 72°C 20 s, finishing with 72°C 5 min, 4°C hold. We routinely setup 2 PCR reactions and pool the volumes before the bead purification.

The PCR products were purified with AMPpure XP beads (Beckman Coulter, Brea, CA, USA), and the concentration after purification was measured on a Qubit instrument (ThermoFischer, Waltham, MA, USA), both procedures performed according to the recommendations of the manufacturer. Before proceeding to

Table 1. Results from 56 clinical prenatal K blood group predictions from pregnant women

Sample No.	Pregnancy[1]	Newborn K serotype	Antenatal result	GA (weeks + days)	Titer	Total reads, *n*	Specific *KEL*01.01* reads	Percentage *KEL*01.01* reads	Note
1	1	K-pos	K-pos	11	1,024	1,010,840	8,296	0.82070	K-pos after re-analysis[2]
2	2a	K-pos	K-pos	10	2,048	26,135,876	1,265,942	4.84369	IUT
3	2b	K-pos	K-pos	12	2,048	25,459,985	878,932	3.45221	IUT
4	3	K-neg	K-neg	32	1	17,615,152	4,222	0.02397	
5	4	nd	K-neg	18	64	13,522,391	2,966	0.02193	
6	5	K-neg	K-neg	28	2,048	20,365,062	3,364	0.01652	IUT
7	6	nd	K-neg	28+3	4	21,527,271	2,486	0.01155	
8	7a	K-neg	K-neg	27	512	14,377,011	2,263	0.01574	
9	7b	K-neg	K-neg	27	512	17,224,605	2,654	0.01541	
10	8	K-pos	K-pos	16	8	18,914,514	1,527,778	8.07728	
11	9	nd	K-neg	23	128	16,232,624	2,025	0.01247	IUT
12	10	K-pos	K-pos	14	128	15,378,883	511,861	3.32834	
13	11	nd	K-pos	13	1,024	18,074,876	176,757	0.97792	IUT
14	12	K-pos	K-pos	22	1	13,915,857	860,228	6.18164	
15	13	K-neg	K-neg	18-19	2	17,011,758	3,171	0.01864	
16	14a	K-neg	K-neg	13+6	4	14,459,623	3,294	0.02278	
17	14b	K-neg	K-neg	20+1	4	17,352,273	3,260	0.01879	
18	15a	K-neg	K-neg	12	2,048	17,221,649	2,800	0.01626	
19	15b	K-neg	K-neg	10	2,048	17,239,164	2,974	0.01725	
20	16	nd	K-neg	17+5	128	15,843,360	3,017	0.01904	Miscarriage
21	17	K-pos	K-pos	16+1	1,024	17,717,110	524,215	2.95881	IUT
22	18	K-neg	K-neg	38+6	1	13,557,559	3,233	0.02385	
23	19	K-pos	K-pos	13	2,048	14,852,063	748,863	5.04215	IUT
24	20a	K-neg	K-neg	17+6	256	16,533,463	3,173	0.01919	
25	20b	K-neg	K-neg	18+4	256	6,759,303	1,391	0.02058	
26	21	K-neg	K-neg	14	16	13,031,945	3,950	0.03031	
27	22a	K-neg	K-neg	14+1	1	18,250,237	3,002	0.01645	
28	22b	K-neg	K-neg	34+6	1	17,941,630	2,965	0.01653	
29	23a	K-neg	K-neg	16	1	15,323,994	3,034	0.01980	
30	23b	K-neg	K-neg	13	1	5,171,393	2,130	0.04119	
31	24	K-neg	K-neg	17	256	14,920,155	2,885	0.01934	
32	25	K-pos	K-pos	13+5	4	18,050,682	618,588	3.42695	
33	26a	nd	K-neg	12+5	64	16,554,791	3,365	0.02033	
34	26b	nd	K-neg	?	64	14,938,969	2,993	0.02003	
35	27a	K-neg	K-neg	14+5	1,024	16,901,268	2,753	0.01629	
36	27b	K-neg	K-neg	17	1,024	16,055,544	9,642	0.06005	
37	27c	K-neg	K-neg	18+5	1,024	3,723,456	1,008	0.02707	
38	28a	nd	K-neg	14	4,000	7,623,476	1,560	0.02046	
39	28b	nd	K-neg	16+2	4,000	4,591,970	1,423	0.03099	
40	29	nd	K-neg	16+6	1	5,549,998	1,613	0.02906	
41	30	K-neg	K-neg	22	?	6,257,514	1,567	0.02504	
42	31	K-pos	K-pos	16+2	16	18,121,034	235,811	1.30131	
43	32	Not born	K-neg	13+3	1	8,879,996	1,613	0.01816	
44	33	K-neg	K-neg	20	2	4,830,241	1,149	0.02379	Miscarriage
45	34	No info	K-neg	16		5,338,870	911	0.01706	
46	35	No info	K-pos	10+5	1,024	15,290,264	609,130	3.98378	Fetus died in utero
47	36	K-pos	K-pos	28	1	3,491,254	129,290	3.70325	
48	37	nd	K-neg	28+3	1	4,977,506	1,168	0.02347	
49	38	Not born	K-neg	16+4	32	5,900,503	1,313	0.02225	
50	39	Not born	K-pos	15+5	2	7,109,332	254,791	3.58390	
51	40	Not born	K-neg	12+1	?	4,561,316	761	0.01668	
52	41a	Not born	K-neg	19+4	16	4,194,359	836	0.01993	
53	41b	Not born	K-neg	15+2	8	4,729,103	774	0.01637	
54	42	Not born	K-neg	16+3	64	4,819,390	1,017	0.02110	
55	44	Not born	K-neg	14+5	1	4,898,913	856	0.01747	
56	45	No info	K-neg	27+1	?	5,342,534	1,034	0.01935	

With two exceptions, IUTs were given to women with K-positive fetuses and an anti-K titer of 1,024 or higher. In sample No. 1 the newborn was anemic and only after re-analysis with the double bioinformatics approach was the fetus predicted to be positive. A few samples were repeat samples from the same pregnancy and always gave the same result. pos, positive; neg, negative; nd, not done; IUT, intrauterine transfusion, ? denotes no information. [1] Repeat samples in the same pregnancy denoted by letter a, b, or c. [2] The anemia was most likely due to the anti-K antibodies.

sequencing, the concentration of purified PCR product was required to be more than 5 μg/mL. Also, the PCR products were tested on a Bioanalyzer (Agilent, Santa Clara, CA, USA) to ascertain if spurious amplification had occurred. More than 90% of the PCR product was required to be in a single peak to proceed with sequencing. If these criteria were not met, a novel amplification was undertaken.

Sequencing of Amplicons

NGS sequencing was performed on an Illumina MiSeq instrument (Illumina, San Diego, CA, USA) and sequencing was done in one direction using the 50-bp kit from Illumina as previously described [13]. The aim for the cluster density was to be between 800 and 1,000/mm^2.

Data Analysis

The FASTQ data of all the clinical samples were routinely analyzed in two ways: using the FastQC software v.0.11.5 (Bioinformatics Group at the Babraham Institute, UK; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and a simple Unix-based grep string search on a Linux formatted PC, using a bash script [15]. The command grep stands for g/re/p, which means: search **g**lobally for **re**gular expressions and **p**rint the result. This grep search command in Linux initiates a search in a document for a particular string of characters and returns certain specified information, for instance the number of times that the character string occurred in the document. The FastQC software analyzed DNA sequences with a frequency of more than 0.1%. The reads with the same sequence were binned, and the number of reads of each allele was counted. The grep search result was the percentage with which the specific *KEL*01.01* string occurred in the FASTQ data relative to the sum of the other strings.

To investigate the effect of varying the string lengths on the analysis, a range of search strings covering progressively shorter parts of the amplicon, both for the *KEL*01.01*-positive and negative sequences, were generated (online suppl. Table 2). Searches with each search string were performed on 3 randomly selected *KEL*01.01*-negative and 3 randomly selected *KEL*01.01*-positive FASTQ data sets from different clinical samples. For routine analysis, the strings in online supplementary Table 3 were used, as previously described [14].

For analysis of clinical samples, a cut-off value of 0.05% was originally chosen. As K-negative samples accumulated, the cut-off was repeatedly recalculated as 5× the standard deviation of the percentage of *KEL*01.01* reads from all K-negative samples (both presumed negative and with postnatal negative serology) plus the average percentage of *KEL*01.01* reads from all the K-negative samples inspired by Lardeux et al. [16]. The reason for the repeated recalculation of the cut-off was the limited empirical experience of the variation of negative samples. One sample had been predicted K negative and the repeat sample 3+ weeks later was found to be marginally above the cut-off (sample 36 with 0.06005%) and reported as inconclusive to the clinicians. We requested a repeat sample, and that was below the cut-off of 0.05%. On the basis of the marginally increased sample, the cut-off was recalculated to 0.0661%. All other samples below this figure were reported as predicted K negative to the clinicians and all samples above this cut-off were reported as predicted as K positive to the clinicians. The reason for the marginally increased *KEL*01.01* reads in sample 36 is unclear.

Reproducibility

To estimate assay reproducibility, three independent determinations on the same pooled *KEL*01.01*-positive cfDNA from a clinical sample were performed. The experiments were run on 3 different days and performed by the same person.

Serology

Anti-K titer and K phenotype were determined as described [17, 18]. The Column technique was used for determining the K type of the mother. It was essential to use flow cytometry for the determination of K type of the newborn and use of flow cytometry for this purpose was mandatory when K-negative transfusions had been given to the fetus. The flow cytometry was performed as described [19].

Data on the K phenotype were obtained for 34 newborns. The anti-K titer was determined at the date of the prenatal sample or in some instances a few days before that date.

Spike-In Experiment

To experimentally assess the sensitivity of the assay, genomic DNA was purified from a whole blood sample from a serologically K-negative (*KEL*02* homozygous) and a serologically heterozygous K/k person. Per PCR reaction 20,000 pg of DNA from the K-negative person was spiked with nominally 0, 10, 25, and 50 copies of K-positive (*KEL*01.01* positive) DNA from a heterozygous individual and examined in the assay. DNA concentrations were measured on a Qubit instrument (ThermoFisher) and diluted independently in molecular biology grade water (Lonza, BioNordika, Copenhagen).

Calculation of the Minimal Number of Molecules to Be Interrogated

For estimating the minimal number of original molecules (n) that must be interrogated to denote a negative result, the following formula was used [11]:

$$n > \ln \alpha / \ln (1 - p) \quad (1)$$

where n is the number of molecules carrying the rs8176058 SNV, α is the probability of not finding a *KEL*01.01* SNV, and p is the fraction of fetal cfDNA.

Results

We set up an NGS-based analysis for non-invasive fetal K genotyping to predict the fetal K status. We investigated additional important design-related and experimental issues relevant to assay accuracy. Q30 was above 90% for all FASTQ sequencing data and in most cases above 95%.

Assay Considerations

An important PCR design issue is how the PCR amplicon size may affect PCR amplification of cfDNA. The relationship between the percentage of cfDNA theoretically amplifiable with amplicon length is shown in online supplementary Figure 1a, and a visualization using experimental data is shown in online supplementary Figure 1b for comparison. Both figures demonstrate that with increasing amplicon length a smaller fraction of cfDNA will be available for amplification. Also, an increased bias towards preferential amplification of maternal cfDNA with increased amplicon length is seen in online supplementary Figure 1. For every 1-bp increase in gene-specific amplicon length, the number of specific cfDNA molecules

available for amplification will decrease by approximately 0.7%. The following calculation further demonstrates the effect of amplicon size. A very small amplicon size would be about 41 bp (two primers of each 20 bases amplifying an interprimer sequence of one base). A gene-specific amplicon of 58 bases will potentially enable amplification of about 59% of fetal cfDNA template molecules carrying the *KEL*01.01* SNV:

$$\text{SNV: } \left[\frac{(143-58)\,\text{bp}}{(143)\,\text{bp}} = 0.59\right]$$

Conversely, for maternal cfDNA, about 65% of cfDNA molecules carrying the *KEL*01.01* SNV will be potentially amplifiable (online suppl. Fig 1a), as a result of a slightly longer average fragment size of 166 bp [14]. To estimate the proportion of cfDNA fragments that are actually amplifiable with a given amplicon length, it is necessary to take into consideration the actual size distribution of cfDNA. The percentage of amplifiable cfDNA fragments was calculated based on experimental data from Prof. Dennis Lo [14], in which the frequency of each fragment length from 20 to 250 bp was estimated for fetal and maternal cfDNA in plasma. The calculation was as follows:

$$\frac{(\text{fragment length} - \text{amplicon length} + 1) \times \text{fragment frequency}}{[(2 \times \text{amplicon length} + 1) + (\text{fragment length} - \text{amplicon length} + 1)]}.$$

This result of the amplifiable fraction was calculated for each fragment length, which was then cumulated for all results equal to and higher than the amplicon length, and finally divided by the total fragment frequency, multiplied by 100 to give the percentage. This was calculated for every 10 bp in the interval from 20 to 130 bp for both fetal and maternal DNA (online suppl. Fig. 1b).

The cumulated risk of genetic variation in the gene-specific part of the primers was 0.022% (data not shown). This value is presumably close to the theoretical overall maximal allele drop-out frequency caused by target sequence variation based on sequence information in the Genome Aggregation Database (gnomAD) database (June 10, 2020). We did not look at the ethnic SNV frequency variation.

For the sequence analysis, the prediction of fetal K status was based on the number of relative reads, i.e., the percentage of *KEL*01.01* reads relative to all non-*KEL*01.01* reads as found by FastQC or grep analysis. The absolute number of reads must be enough to draw a conclusion. In the absence of a control for the presence of fetal cfDNA, we used equation 1 to calculate the theoretical minimum number of cfDNA molecules before amplification that must be examined to be able to diagnose a fetus. By setting $\alpha = 0.01$ and the paternal contribution of the fetal fraction at 1%, it can be calculated from equation 1 as $n > \ln 0.01/\ln(1 - 0.01)$, which equals more than 458 independent cfDNA molecules. This means that the likelihood that none of the 458 molecules are positive for *KEL*01.01* is about 1% despite a fetal fraction of 2%. Under ideal amplification circumstances, i.e., amplification of all amplifiable cfDNA molecules with the same amplification efficiency, the fetal fraction of amplifiable molecules should remain the same after amplification versus before amplification. This means that the same calculation would be valid for the amplified molecules. This is an important implicit assumption for this kind of assay. The increasing delta value of online supplementary Figure 1 indicates that with longer amplicons, there may be an increase in discrepancy of fetal fraction before versus after amplification.

To examine the effect of the length of the search strings, we tested a range of string lengths (online suppl. Fig. 2). A plateau of the ratio between *KEL*02* and *KEL*01.01* reads was seen between 9 and 35 base long strings. In this interval, the length of the strings did not influence results much albeit with a slight tendency towards higher ratios with a longer string length (online suppl. Fig. 2). For our diagnostic assay, we chose to use a string length of 12 bases as this was well within the plateau and gave high ratios between *KEL*02*/*KEL*01.01* reads. Only two bases were amplified so most of the sequence aberrations detected would be due to errors during primer synthesis.

Very similar results from the reproducibility experiment with three independent experiments were obtained. The cluster density was high, and the number of control reads also had a rather similar pattern. The percentage of *KEL*01.01* reads was highly reproducible, with a variance of 1.6% (Table 2). The precise number of *KEL*01.01* SNP-carrying molecules in this sample was unknown but the percentage of positive *KEL*01.01* reads was close to the average (3.7%) of the positive and presumed positive of all the samples.

The cut-off was set at 0.05% until sample No. 36 was analyzed. The inclusion of this sample increased the cut-off to 0.0661% as calculated from the negative clinical samples at that time. The woman was tested two more times (No. 35 and 37) and both samples were clearly below 0.05% and all were deemed negative for *KEL*01.01*. The observed average background of *KEL*01.01* reads was approximately 0.020% from the grep searches from K-negative women pregnant with K-negative fetuses. On the basis of all 42 known and suspected negative samples the cut-off was recalculated at 0.06163%. We decided, however, to continue to use the figure of 0.0661% as the routine cut-off value.

The 0, 10, 25, and 50 copy spike-in genomic DNA results gave background reads and sensitivity in accordance with the clinical results. The result from one of two experiments is shown in Table 3.

Table 2. Assessment of reproducibility

	Experiment No.		
	1	2	3
Q30	94.33	94.83	94.96
Clusters	1,096±35	1,373±114	1,238±152
All reads	21,150,514	16,745,966	21,493,264
*KEL*01.01*, reads	723,812	586,517	801,670
*KEL*02*, reads	18,686,112	15,102,004	19,217,954
control 1 reads	335	685	430
control 2 reads	3,616	1,899	2,498
control 3 reads	17,771	19,512	30,696
control 4 reads	2,260	4,539	2,648
control 5 reads	623	1,087	864
Reads analyzed	19,434,529	15,716,243	20,056,760
% reads not included in the analysis	8.1132	6.1491	6.6835
*KEL*01.01* reads, % of reads analyzed	3.7244	3.7319	3.9970
Percent, variance	1.6074		

The same cfDNA sample material was analyzed on 3 different days by the same person.

Table 3. The grep analysis results of the spike-in experiment with nominally 0, 10, 25, and 50 copies of *KEL*01.01* genomic DNA into a surplus of *KEL*02* genomic DNA are in accordance with the clinical results

	Sequence	Reads	Spike-in copies	%
1	KEL1_tagc, TAGCGTAAATGGACTTCCTTAAACTTTAACCGAATGCTGA	536	None	0.02074
2	KEL1_tgac, TGACGTAAATGGACTTCCTTAAACTTTAACCGAATGCTGA	8,777	10	0.29943
3	KEL1_cagt, CAGTGTAAATGGACTTCCTTAAACTTTAACCGAATGCTGA	20,521	25	0.59396
4	KEL1_atcg, ATCGGTAAATGGACTTCCTTAAACTTTAACCGAATGCTGA	36,262	50	1.06547
1	KEL2_tagc, TAGCGTAAATGGACTTCCTTAAACTTTAACCGAACGCTGA	2,584,058	None	na
2	KEL2_tgac, TGACGTAAATGGACTTCCTTAAACTTTAACCGAACGCTGA	2,931,229	10	na
3	KEL2_cagt, CAGTGTAAATGGACTTCCTTAAACTTTAACCGAACGCTGA	3,454,978	25	na
4	KEL2_atcg, ATCGGTAAATGGACTTCCTTAAACTTTAACCGAACGCTGA	3,403,382	50	na

These grep data are plotted in online supplementary Figure 3. Both the 10, 25, and 50 copies of *KEL*01.01* genomic DNA spike-in are above the cut-off value and consequently scored as positive in this grep analysis. Each tag appended to the primers has at least two base positions with a base that differs from the other tags, ensuring discrimination in a multiplex reaction. na, not appliccable.

With 95% confidence intervals according to the Poisson distribution, the actual copy number intervals were calculated: nominally 10, 25, and 50 copies would give 4–16 copies, 15–35 copies, and 46–64 copies, respectively. The dilutions of the spike-in samples were made independently. When plotted on a graph, the data points were on an extrapolated line with $R^2 = 0.994$ (online suppl. Fig. 3). The amplification efficiency is estimated to be 79–90%.

The fraction of *KEL*01.01* reads for the 10-copy spike-in sample was 0.3% in the grep analysis, clearly above the cut-off level. The FastQC data analysis detected *KEL*01.01* reads for the 25- and 50-copy spike-in but not for the 10-copy spike-in (online suppl. Table 4), exemplifying the value of the dual data analysis approach. The negative sample gave 0.0207% *KEL*01.01* background reads.

The percentage of *KEL*01.01* from the positive clinical samples were often at a few percent or more, so by extrapolating linearly from a signal from the 50 spike-in copies of 1% means that often a clinical sample will likely contain several hundreds of *KEL*01.01* amplifiable copies per 16 µL of eluate from purification on a Symphony from 4 mL of plasma from a woman with a *KEL*01.01*-positive fetus.

Clinical Prediction of Fetal KEL

We analyzed a total of 56 samples from 43 women during 45 pregnancies. Three pregnancies ended in miscarriage. The samples were from gestational weeks 10–38. An overview of the clinical results from K predictions is shown in Table 1.

Early in the process of development of the assay, we analyzed data by alignment to genomic *KEL*01.01* sequences; however, one sample was assigned as negative whereas postnatal serology showed that the newborn was K positive. This prompted us to amend the data analysis into the described routine procedure and after re-analysis

using that amended procedure, the sample was positive for *KEL*01.01*, albeit with a low percentage of *KEL*01.01* reads (Table 1, sample No. 1). Using the dual analysis, all samples with postnatal serology were in accordance with the prenatal prediction (n = 34).

Three other samples were re-run due to technical problems with the initial sequencing reaction that yielded no useful data and are not included in Table 1. Eight women had not yet given birth at the time of submission of the manuscript, and there was no available postnatal serology for an additional 14 samples. All repeat samples from the same pregnancy gave consistent results.

There were 14 predicted K-positive samples and 42 predicted K-negative samples. In no case was the result of the FastQC analysis and the grep search analysis result incompatible. The grep search result was important in the determination of the negative samples. In all instances of the K-negative fetuses, the fraction of *KEL*01.01* reads was below 0.0661%.

In online supplementary Figure 4 an overview of the results of the clinical samples is shown, sorted after *KEL*01.01* read percentage. The negative sample with the highest number of specific reads had 0.0601% *KEL*01.01* reads, and postnatal serology showed the baby to be K negative. The positive sample with the lowest number of specific reads had 0.8207% *KEL*01.01* reads with the newborn serologically verified. So, there was almost a 14-fold difference between the highest negative sample and the lowest positive sample in this cohort.

Discussion

We have presented data from 7 years of experience from predicting the fetal K status in anti-K-immunized pregnant women. During the very first period of testing, when data were analyzed by an alignment approach, one false-negative sample (Table 1, No. 1) was reported. After this incident, we switched to an alternative analytical approach using both FastQC and the grep search with the search strings in online supplementary Table 2. After re-analysis with this method, the false-negative sample was positive. All other results were in concordance with the known newborn K status.

The development of the approach described in this paper was prompted by an unmet clinical need for early prenatal prediction of fetal blood groups in alloimmunized pregnant women. Over the years our approach has undergone some modifications, most importantly in the data analysis.

To make a robust and streamlined procedure, we used adaptor-appended PCR primers without allele specificity and NGS to enable interrogation of a large number of molecules and counted the number of times that the SNV encoding either the K or k antigen was detected. The primers were designed to produce a small amplicon to interrogate as large a fraction of cfDNA as possible with the purpose of increasing the sensitivity (online suppl. Fig 1a, b).

Streck tubes were used for blood draws to maintain the relative fractions of cfDNA of maternal and fetal origin, respectively, reducing the risk of a false negative outcome. The PCR amplification was done with the Phusion DNA polymerase, a proof-reading DNA polymerase with a low error incorporation rate [20]. This choice was made to keep background reads low.

*KEL*01.01* reads (range 761–9,642, median 2,570) were found in amplicons from PCR-amplified cfDNA in plasma from K-negative women with a K-negative fetus. These background reads were likely introduced during the amplification process. With 40 cycles of amplification and 2 bases amplified with a fidelity 4.4×10^{-7} for the Phusion polymerase, an estimated 0.00352% of the two amplified and sequenced bases will have an error introduced during the amplification process. We observed about 5 times higher background *KEL*01.01* reads, i.e., T bases at this position, in the negative samples (Table 1).

The reproducibility experiment gave a variance of 1.6% based on 3 experiments (Table 2). Indeed, several estimated parameters were quite similar, indicating a highly reproducible analysis. This result was obtained from a sample where the number of *KEL*01.01* molecules per milliliter was unknown but contained between 3.4 and 4.8% *KEL*01.01* reads as analyzed by the described assay.

In the overall design of the assay, we chose not to use a control reaction for the presence of fetal cfDNA. To be useful, a fetal control PCR should closely mimic the amplification efficiency of the primers for the analytical target, and this is not a trivial design problem [21]. However, a good fetal control would be desirable and for instance also help to detect sample mix up with samples from non-pregnant individuals. A negative sample before GA15 is repeated 3 or more weeks later. The cost and resource consumption of a repeat analysis is the same as the first analysis and likely when more experience has been gained, especially with early samples, repeat testing can be omitted. If the first test is taken around GA 10–12 weeks, there is time for a repeat test after GA 15 weeks and instituting IVIG treatment can be started after the first positive test.

In the absence of a fetal control, calculation of the minimal number of cfDNA fragments needed to be interrogated to establish a diagnosis is important to consider. We sequenced at least 1,000-fold more molecules than the estimated minimum number of cfDNA fragments required as calculated from equation 1. With the current approach, however, we cannot empirically determine the number of cfDNA molecules from maternal

plasma with the *KEL*01.01/KEL*02* SNV that are amplified and interrogated.

A very important factor in obtaining a high sensitivity is to probe a large volume of plasma. We used a high plasma equivalent of (16/60) × 4 = 1.07 mL per PCR and pooled the volume from two PCR tubes before purification of the amplified DNA. It was assumed that cfDNA purification of fetal and maternal cfDNA was unbiased. Based on the data from the negative samples, we arrived at a cut-off value of 0.0661%. Thus, the assay should be able to nominally detect one or more cfDNA molecules of *KEL*01.01* among 1,512 (1/0.000661) cfDNA molecules carrying the rs8176058 SNV. At early GA, there are usually no more than 1,500 genome equivalents per milliliter in maternal plasma [22]. The PCR amplification of maternal plasma cfDNA amplifies a slightly smaller fraction of fetally derived cfDNA, this difference is denoted the delta value in online supplementary Figure 1a and b; apart from this, however, the ratio of fetal:maternal cfDNA retains the same relative proportion after amplification as before amplification based on the reasonable assumption that the same amplification efficiency applies to both fetal and maternal cfDNA with the rs8176058 SNV. Sequencing up to 20 million molecules are possible on the MiSeq instrument, and this number should amply compensate for any deviation from ideal amplification conditions; even multiplexing for example, 5 other amplicons in the same sequencing reaction should not impact the prediction precision adversely.

Taken together, these calculations indicate that the assay potentially can detect about 3–5 molecules of *KEL*01.01*-positive SNV in one mL of maternal plasma. Based on the present 34 samples with a postnatal serological determination of the newborn's K phenotype and accepting the result from the dual data analysis as valid comparison, the assay has an accuracy of 100% (95% CI 89.72–100%) and hence a specificity, sensitivity, as well as a positive and negative predictive value of 100%. We obviously do not expect that this figure will remain at this level as more samples are analyzed; however, based on the preliminary data in Table 1, the antenatal assay for noninvasive K prediction can be characterized as promising. The empirical measurement of the spike-in sensitivity agrees well with this notion. A more rigorous determination of the limit of detection and possibly the limit of quantification requires further investigation.

The spike-in experiment, where heterozygous *KEL*01.01* genomic DNA was spiked into 20,000 pg of *KEL*01.01*-negative genomic DNA (approx. 3,000 genome equivalents), showed that the lowest spike-in concentration investigated of 10 copies could be detected and that the amplification efficiency was acceptable. Although experiments using genomic DNA cannot be directly compared with results from cfDNA, the results do indicate a very high sensitivity of the assay where 10 copies were detected within the formally defined cut-off in the grep analysis. The background of 0.02% of the 0-copy spike-in sample (online suppl. Table 4) and the average background of the clinical samples of approximately 0.02% is comparable to results of 0.03% from another group [9].

The two elements of the dual data analyses are mutually supportive: the FastQC analysis interrogates the fastq PF (passing filter) data, and the grep search interrogates all sequences. Improvements of the MiSeq data could perhaps be obtained using patterned flow cells [23] not yet available for the MiSeq.

It has been recommended that a fetal fraction of ≥4% is required to avoid false negatives due to low numbers of the paternal allele [24]. This caution is sound advice; nevertheless, early diagnosis is of paramount importance because the fetus may die in GA 18 weeks in cases with a high titer anti-K antibody unless treatment had been instituted. Often the fetal fraction is below 4% in early pregnancy [25, 26]. Additional testing of more early samples is needed to thoroughly establish the validity of our claim of early prediction. As a precaution, we currently recommend that a negative sample taken before GA 15 weeks is analyzed again on a new sample taken 3 weeks after the first sample. Sending a repeat sample is a decision left to the clinician.

Two *KEL*01.01*-positive samples from GA 11 and 16 + 2 weeks, respectively, had a calculated fetal fraction of 1.6 and 2.6%, both were positive using the current data analysis. We estimated the allele drop-out risk due to genetic variation in the gene-specific primer target sequence at 0.022%, which we consider a low number, which should not preclude the use of the assay.

No indication of increased *KEL*01.01* background was found in a *KEL*01.01*-negative sample that was run on the MiSeq directly after a *KEL*01.01*-positive sample. This is in accordance with the experience from prenatal ABO analysis which has very low background reads [14].

Using the current data analysis, all samples from this cohort that were associated with information of postnatal K status were predicted correctly with a big margin between negative and positive samples. The price for the analysis can be significantly reduced for the individual sample by multiplexing. Our approach to NGS data analysis using FastQC and grep search is very simple and has worked well.

From a clinical perspective, there is also great value in detecting *KEL*01.01*-negative fetuses as this would obviate many parental worries, clinical controls, and ultrasonographic scans throughout the pregnancy. In 100% of the results of a *KEL*01.01* prediction the results have been useful to the clinicians. In case of a predicted *KEL*01.01*-negative fetus the woman should appear in week 25 and

32 for ultrasonographic scans, titer determination, and antibody screen for the occurrence of additional antibody specificities [27]. Depending on specific high-risk circumstances, an extra ultrasonographic scan and titer determination can be done around GA 16–18 weeks.

The predictions of a K-positive fetus were helpful in alerting clinicians to ensure close monitoring of potentially complicated pregnancies. Other factors, such as anti-K titer changes, history of previous births, and ultrasonographic scans, were also included when advising on the management of pregnancies of anti-K immunized mothers. Ultrasonographic scans can usually not reliably be employed before GA 16 weeks and early K prediction were in some cases helpful in the decision of, for instance, early institution of IVIG. Therefore, the predictions of fetal K status had immediate clinical consequences, especially in pregnancies where high titer anti-K was discovered or previously known. We treated the relevant pregnant women with IVIG and IUT. Others have used a regime with rituximab, IVIG, and immunoadsorption, and have shown that no rituximab is detectable in the newborn [28]. This treatment has the advantage that it can be instituted before IUT is possible, but may not be as efficient as IUT and have temporary adverse effects on the maternal immune system.

In conclusion, we have established a robust and sensitive assay with high reproducibility for non-invasive prediction of fetal K status in a clinical setting. We routinely analyze samples from GA 10 weeks. Early diagnosis is necessary to timely and rationally give information and institute intervention and IUT as lifesaving therapy. Important parameters in assay design are a short amplicon and a large plasma equivalent.

Acknowledgments

We thank Birgitte Suhr Bundgaard and colleagues for excellent laboratory work and Prof. Dennis Lo for the generous donation of experimental data underlying online supplementary Figure 1b. Frank Eriksson is thanked for suggestions concerning the statistical considerations.

Statement of Ethics

The reported data are the results from a routine analysis or previously collected anonymized samples used for developing a clinically applicable method. Thus, this report is not concerned with research as defined in Danish law on ethics, and approval by the ethics committee is not required. Permission to obtain postnatal blood group information was obtained from the data holding institutions that kept this information following Danish data regulations.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

The author(s) received no specific funding for this work, which is based on a collection of consecutive prenatal *KEL*01.01* routine analysis results from 2013 to 2021 at Rigshospitalet, Copenhagen, Denmark.

Author Contributions

K.R. conceived the idea for the analytical principle of the analysis and data interpretation, designed primers for the prenatal assay, established the analysis in the laboratory, and drafted the manuscript. F.B.C. contributed to the design of the work and the data interpretation, and substantially contributed to writing the manuscript. M.H.D. supervised the work and promoted the use of the technology among clinicians. All authors contributed to data acquisition and to the manuscript, and approved the final version of the manuscript and agreed to be accountable for all aspects of the work, ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved.

Data Availability Statement

All relevant results from data generated or analyzed during this study are included in this article. Further enquiries can be directed to the corresponding author.

References

1 Castleman JS, Moise KJ Jr., Kilby MD. Medical therapy to attenuate fetal anaemia in severe maternal red cell alloimmunisation. Br J Haematol. 2021;192(3):425–32.

2 Slootweg YM, Lindenburg IT, Koelewijn JM, Van Kamp IL, Oepkes D, De Haas M. Predicting anti-Kell-mediated hemolytic disease of the fetus and newborn: diagnostic accuracy of laboratory management. Am J Obstet Gynecol. 2018;219(4):393–e8.

3 Zwiers C, van der Bom JG, van Kamp IL, van Geloven N, Lopriore E, Smoleniec J, et al. Postponing early intrauterine transfusion with intravenous immunoglobulin treatment; the PETIT study on severe hemolytic disease of the fetus and newborn. Am J Obstet Gynecol. 2018;219(3):291–e9.

4 Maisonneuve E, Dugas A, Friszer S, Toly-Ndour C, Cariot L, Dhombres F, et al. Effect of intravenous immunoglobulins to postpone the gestational age of first intrauterine transfusion in very severe red blood cell alloimmunization: a case-control study. J Gynecol Obstet Hum Reprod. 2021;50(7):102119.

5 Rieneck K, Bak M, Jonson L, Clausen FB, Krog GR, Tommerup N, et al. Next-generation sequencing: proof of concept for antenatal prediction of the fetal Kell blood group phenotype from cell-free fetal DNA in maternal plasma. Transfusion. 2013;53(11 Suppl 2):2892–8.

6 van der Schoot CE, Winkelhorst D, Clausen FB. Noninvasive fetal blood group typing. In: Page-Christiaens L, Klein H, editors. Noninvasive prenatal testing (NIPT). Applied genomics in prenatal screening and prenatal diagnosis. Amsterdam: Elsevier; 2018. p. 125–56.

7 Finning K, Martin P, Summers J, Daniels G. Fetal genotyping for the K (Kell) and Rh C, c, and E blood groups on cell-free fetal DNA in maternal plasma. Transfusion. 2007;47(11):2126–33.
8 O'Brien H, Hyland C, Schoeman E, Flower R, Daly J, Gardener G. Non-invasive prenatal testing (NIPT) for fetal Kell, Duffy and Rh blood group antigen prediction in alloimmunised pregnant women: power of droplet digital PCR. Br J Haematol. 2020;189(3):e90–e4.
9 Orzinska A, Guz K, Mikula M, Kluska A, Balabas A, Ostrowski J, et al. Prediction of fetal blood group and platelet antigens from maternal plasma using next-generation sequencing. Transfusion. 2019;59(3):1102–7.
10 Rieneck K, Clausen FB, Dziegiel MH. Next-generation sequencing for antenatal prediction of KEL1 blood group status. Methods Mol Biol. 2015;1310:115–21.
11 Rieneck K, Clausen FB, Dziegiel MH. Noninvasive antenatal determination of fetal blood group using next-generation sequencing. In: Bianchi D, Norwitz ER, editors. Molecular approaches to reproductive and newborn medicine. New York: Cold Spring Harbor; 2015. p. 315–23.
12 Chitty LS, Lo YMD. Noninvasive prenatal screening for genetic diseases using massively parallel sequencing of maternal plasma DNA. In: Bianchi DW, Norwitz ER, editors. Molecular approaches to reproductive and newborn medicine. New York: Cold Spring Harbor; 2015. p. 295–314.
13 Tsui NB, Jiang P, Chow KC, Su X, Leung TY, Sun H, et al. High resolution size analysis of fetal DNA in the urine of pregnant women by paired-end massively parallel sequencing. PLoS One. 2012;7(10):e48319.
14 Rieneck K, Egeberg Hother C, Clausen FB, Jakobsen MA, Bergholt T, Hellmuth E, et al. Next generation sequencing-based fetal ABO blood group prediction by analysis of cell-free DNA from maternal plasma. Transfus Med Hemother. 2020;47(1):45–53.
15 Ramey C. Bash reference manual, version 5.0. Case Western Reserve University: Brian Fox, Free Software Foundation Inc.; 2019.
16 Lardeux F, Torrico G, Aliaga C. Calculation of the ELISA's cut-off based on the change-point analysis method for detection of *Trypanosoma cruzi* infection in Bolivian dogs in the absence of controls. Mem Inst Oswaldo Cruz. 2016;111(8):501–4.
17 Lapierre Y, Rigal D, Adam J, Josef D, Meyer F, Greber S, et al. The gel test: a new way to detect red cell antigen-antibody reactions. Transfusion. 1990;30(2):109–13.
18 Nordvall M, Dziegiel M, Hegaard HK, Bidstrup M, Jonsbo F, Christensen B, et al. Red blood cell antibodies in pregnancy and their clinical consequences: synergistic effects of multiple specificities. Transfusion. 2009;49(10):2070–5.
19 Dziegiel MH, Hansen MH, Haedersdal S, Barrett AN, Rieneck K, Main KM, et al. Blood chimerism in dizygotic monochorionic twins during 5 years observation. Am J Transplant. 2017;17(10):2728–32.
20 Li M, Diehl F, Dressman D, Vogelstein B, Kinzler KW. BEAMing up for detection and quantification of rare sequence variants. Nat Methods. 2006 Feb;3(2):95–7.
21 Delghandi M, Delghandi MP, Goddard S. The significance of PCR primer design in genetic diversity studies: exemplified by recent research into the genetic structure of marine species. Methods Mol Biol. 2022;2392:3–15.
22 Jensen MPS, Damkjaer MB, Clausen FB, Ali HA, Hare KJ, Dziegiel MH, et al. Targeted Rhesus immunoglobulin for RhD-negative women undergoing an induced abortion: a clinical pilot study. Acta Obstet Gynecol Scand. 2019;98(9):1164–71.
23 Singer GAC, Fahner NA, Barnes JG, McCarthy A, Hajibabaei M. Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. Sci Rep. 2019;9(1):5991.
24 Wienzek-Lischka S, Bachmann S, Froehner V, Bein G. Potential of next-generation sequencing in noninvasive fetal molecular blood group genotyping. Transfus Med Hemother. 2020;47(1):14–22.
25 Ashoor G, Syngelaki A, Poon LC, Rezende JC, Nicolaides KH. Fetal fraction in maternal plasma cell-free DNA at 11-13 weeks' gestation: relation to maternal and fetal characteristics. Ultrasound Obstet Gynecol. 2013;41(1):26–32.
26 Wang E, Batey A, Struble C, Musci T, Song K, Oliphant A. Gestational age and maternal weight effects on fetal cell-free DNA in maternal plasma. Prenat Diagn. 2013;33(7):662–6.
27 Dziegiel MH, Krog GR, Hansen AT, Olsen M, Lausen B, Nørgaard LN, et al. Laboratory monitoring of mother, fetus, and newborn in hemolytic disease of fetus and newborn. Transfus Med Hemother. 2021 Sep 8;48(5):306–15.
28 Capraru A, Müller M, Daskalakis M, Bulgheroni LR, Hustinx HC, Zeerleder SS, et al., editors. Severe anti-K alloimmunisation during pregnancy: personalized treatment with immunoadsorption, intravenous immunoglobulin and rituximab. Transfus Med Hemother 2020;47:22.