

Unbiased Next Generation Sequencing-based Approach for Minimal Residual Disease Monitoring in Multiple Myeloma

مانفرد كوبر^{1,2} و رنا الحلاق³

Manfred Kuepper^{1,2} and Rana Hallak³

¹قسم الكيمياء الحيوية والأحياء الدقيقة، كلية الصيدلة، الجامعة العربية الدولية، سورية، ²معهد المناعة وعلم الوراثة، كايسرزلوترن،

ألمانيا، ³قسم الكيمياء الحيوية والأحياء الدقيقة، كلية الصيدلة، جامعة الرشيد الخاصة، سورية

¹Arab International University, Faculty of Pharmacy, Dept. of Biochemistry, Ghabareb, SAR, ²Institute for Immunology and Genetics, Kaiserslautern, Germany, ³Al Rasheed University, Faculty of Dentistry, Ghabareb, SAR, ⁴Rana Hallak Clinical Laboratory, Damascus Kafar Sousseh, SAR

المخلص Abstract:

يتصف الورم النقوي MM بتراكم الخلايا البلازمية الخبيثة MMPC في نخاع العظم. ورغم اللجوء حالياً لتطبيق المعالجة عالية الجرعة HDT، وإدخال مشبطات البروتوزوم والعوامل المعدلة للمناعة فإن حدوث النكس بعد تطبيق العلاج عالي الجرعة، خلال مدة لا تتجاوز العامين وسطياً، يبقى أمر متوقع. هذا وقد أصبح التفاعل التسلسلي الكمي المعروف بـ qASO-PCR، والذي يستهدف الموقع معاد الترتيب من سلاسل الغلوبولين المناعي الثقيلة IgH، من المقاربات المعتمدة في رصد المرض المتبقي الأدنى MRD. ورغم ذلك فإنه لا بد من اللجوء لتصميم بادئات فردية للـ PCR لكل مريض على حدة، مما يجعل الطريقة غير ملائمة للمخابر السريرية المعيارية، ذلك أن هذا الأمر يتطلب خبرة كبيرة ومعرفة واسعة بعلم الوراثة المناعية. وأكثر من هذا وذاك، فإن النسائل الورمية عرضة للإصابة بطفرات إضافية (ضمن منطقة CDR3 مثلاً) يصعب تمييزها.

في هذا البحث سنتطرق لتقنية التسلسل المعروفة بـ NGS massive-parallel ultra-deep sequencing والمرتكزة على مجموعات بادئة تتيح التحري الفعال لأي من الأضداد معادة الترتيب ومورثات السلاسل الخفيفة MMPCs. ومن خلال استخدام الطريقة مزدوجة السوية للتقييم الكمي للـ IgH في المستوى الأول، والتقييم الكمي لسلاسل كبا ولابدأ الخفيفة في المستوى الثاني، فإننا نتمكن من التغلب على ظاهرة إغفال أو تجاوز المرضى ذوي تسلسلات IgH اللامضخمة.

هذا وإن استخدام بروتوكول PCR معياري موحد لكل مريض يجعل الطريقة قابلة للتطبيق في أي مخبر جزيئي سريري، بل ويسمح بالتحري الكمي لكل تسلسل نسيلي بغض النظر عن التغيرات التي تطرأ على تسلسل الدنا بسبب الطفرات المستمرة، مما يمنح هذه الطريقة موثوقية ومعالجة عالية مقارنة بالـ qASO-PCR.

إن هذا التطبيق عالي الحساسية للتقييم الكمي لكل من الـ IgH والسلاسل الخفيفة بحيث يمكن من تحري خلية ورمية واحدة من أصل 33.000 خلية بائية سليمة في تفاعل PCR وحيد. وقد لوحظت هذه الحساسية في 5 حالات من الميولوما وفي نسائل خلوية للمفوما بوركت معادة الترتيب من VH1 لـ VH5 وفي خطوط خلوية لحالتي ميولوما معادة الترتيب لسلاسل كبا ولابدأ الخفيفة على التوالي. وستعتمد هذه الطريقة كتقنية متعددة الاستعمالات تسمح بمسح MRD في أورام خلايا B الأخرى كالـ B-CLL والتي يمثل فرط الطفرات الجسدية فيها صعوبة وعقلة لتطبيق الـ qASO-PCR.

Multiple myeloma (MM) is characterized by the accumulation of malignant plasma cells (MMPC) in the bone marrow. Despite the implementation of high-dose therapy (HDT) and the introduction of proteasome inhibitors and immunomodulatory agents, relapse in MM still occurs at a median of 2 years post-HDT. Quantitative allele-specific oligonucleotide polymerase chain reaction (qASO-PCR) targeting the rearranged immunoglobulin heavy chain locus (IgH) is a well-established monitoring approach for minimal residual disease (MRD). However, for each

patient individual PCR primers and assays have to be designed and evaluated, rendering the method unsuitable for standard clinical laboratories since this requires detailed knowledge of immunogenetics and vast experience. Moreover, tumor clones are prone to undergo further mutations, for instance within the CDR3 region, by which the tumor clone – or a subclone – possibly evades detection.

Here we present a targeted next-generation sequencing (NGS) massive-parallel ultra-deep sequencing approach based on consensus and family primer sets that enable the effective and unbiased detection of any rearranged IgH and light chain genes of MMPCs. By using a two-tiered approach of IgH quantification in the first line and kappa and lambda light chain quantification in the second line, we are able to circumvent most drop outs of patients with non-amplifiable IgH sequences. Using one standardized PCR protocol for every patient makes the method applicable to any clinical molecular laboratory. Moreover, it allows quantitative detection of any clonal sequence regardless of changes in the DNA sequence due to ongoing mutations, making this approach more robust and more reliable than qASO-PCR.

The presented approach is highly sensitive for both, IgH and light chain quantification, detecting 1 tumor cell among 330.000 normal B cells out of one PCR reaction (6pg target DNA among 2µg polyclonal background DNA). The sensitivity has been demonstrated in 5 different Myeloma and Burkitt Lymphoma cell lines rearranged for VH1 through VH5 and 2 Myeloma cell lines rearranged for kappa and lambda light chain, respectively.

As a versatile technique, this approach also allows for MRD screening in other B-cell neoplasia, such as B-CLL, a malignancy, in which ongoing somatic hypermutations can pose an obstacle to qASO-PCR.

:Key words الكلمات المفتاح

NGS، الورم النقوي المتعدد، B-NHL، ضبط هدأة المرض، qASO-PCR، PCR الغلوبولين المناعي

Next-generation sequencing, multiple myeloma, B-NHL, remission control, MRD screening, qASO-PCR, immunoglobulin PCR

Introduction

Multiple myeloma (MM) is an end point in the spectrum of plasma cell (PC) dyscrasias. Proliferating malignant plasma cells (MMPCs) damage target organs, such as bones, the bone marrow and kidneys (1). With approx. 15.000 new cases per year in the US, MM is the second most common hematologic malignancy (2). In spite of the advancement of new targeted therapies, like proteasome inhibitors and immunomodulatory agents, MM remains incurable, with a median survival time of 3 to 4 years (3). Relapse in MM still occurs at a median of 2 years post-HDT.

Autologous stem cell transplantation has become an important therapeutic approach since graft-versus-host disease is not a risk and elder patients can tolerate the procedure, while allogeneic bone marrow transplantation (BMT) is of limited use

with respect to the median age of 64 years at diagnosis and significant transplant-related complications. A number of studies however, demonstrated that the blood stem cell harvest products from patients with MM are frequently contaminated with malignant cells (4) - (8), which alone makes a quantitative method for MRD monitoring mandatory.

The amount of circulating plasma cells is related to a shorter over-all survival, while the load of monoclonal plasma cells in the stem cell harvest is associated with a shorter relapse-free survival (2), (9), (10). Moreover, the kinetics of the BM tumor load in course of HDT has been demonstrated as a predictive parameter (11), just like the post-transplantation BM tumor load (12).

For years, multiparametric flow cytometry (MFC), a fast, simple and cost-effective method, was gold standard for remission control and MRD screening in MM (13). Anyhow, the limited sensitivity of 10^{-3} to 10^{-4} (2), (14), (15),

is suitable for relapse detection but does not allow detection of its onset, which would provide greater therapeutic success. As well, owed to this limited sensitivity, remission control for MM in particular would require bone marrow biopsy at the location of the former disease focus; it goes without saying that this procedure cannot be repeated in short intervals and stands against patients' compliance.

Quantification of clonotypic IgH VDJ rearrangements by qASO-PCR improved the sensitivity to 10^{-5} , has become a well-established approach for MRD monitoring (10), (11), (15). However, patient-specific PCR assays have to be individually designed and evaluated, requiring special skills and rendering the method unfit for standard clinical molecular laboratories. Moreover, in qASO-PCR a considerable proportion of heavy-chain amplification fails, in part due to ongoing mutations of the tumor clone within the highly susceptible primer binding site (16).

Martinez-Lopez et al. (17) compared the prognostic value of next-generation deep sequencing using the commercial LymphoSIGHT™ method (Sequentia, Inc., San Francisco, CA), vs. qASO-PCR and MFC in MM patients with complete response (CR) or very good partial response (VGPR) after front-line therapy, respectively. They demonstrated that patients who appeared MRD+ by MFC but MRD- by sequencing had a significantly longer time to tumor progression (TTP) compared to patients who were MRD+ by sequencing, hinting to possible artifacts that MFC may produce. More importantly, the deep sequencing approach identified a tumor marker in 91% of MM patients in contrast to only 70% with qASO-PCR, demonstrating significant superiority of the former methodology.

Yao et al. (18) and Ha et al. (19) demonstrated superiority of NGS-based deep sequencing over qASO-PCR and fragment analysis, respectively using the LymphoTrack IgH and IgK assays (Invivoscribe Inc., San Diego, CA, USA).

Our goal was to develop a non-proprietary assay that is suitable for routine application by eliminating the drawbacks of qASO-PCR while maintaining the high sensitivity and specificity of a PCR reaction. A targeted NGS deep sequencing approach based on consensus and

family primer sets allows the effective and highly sensitive detection of rearranged *IgH*- and light chain genes of MMPC. A two-tiered approach of IgH quantification in the first line and multiplexed kappa (IgKL) and lambda light chain (IgLL) quantification in the second line circumvents most drop outs of patients with non-amplifiable IgH sequences.

The *IgH* gene locus on chromosome 14 (14q32.3) includes 46-52 functional and 30 non-functional variable (V_H) gene segments, containing three conserved framework (FR) and two variable complementarity-determining regions (CDRs), 27 functional diversity (D_H) gene segments, and 6 functional joining (J_H) gene segments. During B-cell development these segments are rearranged generating a unique $V_H - D_H - J_H$ combination.

Any rearranged IgH sequences in the patient's sample are amplified using family-specific V_H FR1 forward primer and consensus J_H reverse primer. However, the rearranged V_H segment is typically subject to somatic hypermutations, which can impede DNA amplification if they occur within the primer binding sites of the involved V_H gene segment. In case of such an allele drop-out, the light chains are amplified alternatively. The PCR product is subjected to massive parallel sequencing on an NGS platform and clonotypic sequences are identified among all specific reads by means of bioinformatics. As a result, the number of clonotypic reads in proportion to all amplified immunoglobulin sequences is assessed. Using a reference DNA as internal control allows determining the copy number of the clonotypic DNA in the sample, thus assessing the number of residual tumor cells within one order of magnitude.

The use of family-specific forward primer binding to the FR1 region eliminates the need for patient-specific primer, rendering the method convenient for diagnostic routine while making it less susceptible for PCR drop-outs as consequence of ongoing mutations within the CDR3 region.

Materials and Methods

Cell Lines and Patient DNA

We used the commercially available myeloma cell lines OPM-2 (IgG Lambda; DSMZ-No. ACC-50)(20), U266 (IgE Lambda; ATCC No.

TIB-196)(21), and NCI (IgA Kappa; ATCC No. CRL-9068)(22) which have been kindly provided by the myeloma laboratory of the Otto Meyerhoff Center, Int. Med. V, University Hospital Heidelberg, Germany. The laboratory of Prof. Dr. M. Pfreundschuh, José Carreras Center, Int. Med. I, University Hospital Homburg/Saar, Germany, kindly provided the Burkitt lymphoma cell lines Raji (DSMZ-No. ACC-319)(23), BL29(24), and CA46 (DSMZ-No. ACC-73)(25), (26), as well as the lymphoblastoid cell lines (LCLs) G57, and P69, which had been established from myeloma cells isolated from patient samples (Table 1).

DNA that was isolated from washed cell pellets of healthy donors obtained from buffy coat served as polyclonal control.

Table 1: Cell lines used in the study

Type	Name	IgH	IgL
myeloma	OMP2	n.a.	IgLL
	U266	VH1	IgLL
	NCI	n.a.	IgKL
Burkitt lymphoma	Raji	VH3	n.a.
	BL29	VH4	n.a.
	CA46	VH5	n.a.
LCL	G57	VH3	IgLL
	P69	VH2	IgLL

DNA Isolation

DNA was isolated from washed cell pellets using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). In brief, the cell pellets were thawed

from -80°C at room temperature (RT); dry pellets were resuspended in 200µl PBS and processed according to the protocol of the manufacturer.

DNA concentration and purity were determined by UV spectroscopy at 260nm and 280nm, respectively, using a BioSpectrometer Basic (Eppendorf, Hamburg, Germany). Concentration of all preparations was adjusted to 100 ng/µl using pure water.

PCR Primer

Multiplex amplification of *IgH* VH allele families is achieved using family-specific primer within the FR1 region and two JH consensus primer from the intronic region, as published elsewhere (Küpper, et al., 1996) (Table 1), yielding PCR products of approximately 350bp. This primer set proved to be superior to the initially tested unified consensus primer panel as suggested by the European BIOMED-2 collaborative study (27) which suffered from inferior sensitivity and biased amplification efficiency of certain VH families (data not shown).

For the *IgLK* and *IgLL* genes, consensus primer were designed from the FR1- and J-regions (Table 2 and Table 3), yielding PCR products of 282 to 315bp for kappa and 269 to 380bp for lambda, respectively. All primer have been optimized for identical annealing temperatures, to facilitate multiplexing of *IgLK* and *IgLL*.

All primer were ordered from Biolegio (Nijmegen, the Netherlands) and HPLC purified.

Table 1: IgH FR1 and JH primer sequences

Oligonucleotide	Sequence
V _H 1	CTCAGTGAAGGTTTCCTG
V _H 2	GTCCTGCGCTGGTGAAAC
V _H 3	GGGTCCCTGAGACTCTCC
V _H 4	CCTGTCCCTCACCTGCGC
V _H 5	AAAAAGCCCGGGGAGTCT
V _H 6	CTGTGCCATCTCCGGGGA
J _H intron 3,6	ACTTACCTGAGGAGACGG
J _H intron 1,2,4,5	GACTCACCTGAGGAGACG

Table 2: IgLK and JK primer sequences

Oligonucleotide	Sequence
V _K 1a	GGAGACAGAGTCACCATCACTTG
V _K 2a	CTCCATCTCCTGCAGGTCTAG
V _K 3a	CTGCAGGGCCAGTCAGAG
V _K 5a	CTTTGTCTCCTGGAGTCGCTG
J _K 4	GTTTGATCTCCACCTTGGTCCC
J _K 5	TCGAGTCCCTTGGCCGAAGGTG

Table 3: IgLL and JL primer sequences

Oligonucleotide	Sequence
V _L 12a	GGTCCTGGGCTCAGTCTG
V _L 3c	TGGTACCAGCAGAAGCCAGG
V _L 4a	CCAGCCTGTGCTGACTCA
V _L 7a	CAGACTGTGGTGACTCAGGAG
J _L HD	CTAGGACGGTGAGCTTGGTCCC
J _L 5	AGACTCATCTAGGACGGTCAG

PCR Conditions

Multiplex amplification of rearranged *IgH*-genes is carried out in a final volume of 50 µl with 20 pmol of each primer (Biolegio, Nijmegen, the Netherlands), 10 nmol of each deoxyribonucleoside triphosphate (Eurogentec, Liege, Belgium), 2mM MgCl₂, and 1.25 U AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA) on a Biometra Professional Thermocycler (Biometra, Göttingen, Germany). The *IgH*-PCR cycling conditions are as follows: initial denaturation 5 min at 95°C; 50 cycles of 94°C for 30s, 60°C for 30s, and 72°C for 45s; final extension of 10 min.

The *IgLK / IgLL-multiplex PCR* is carried out using the Qiagen Multiplex PCR Kit (Qiagen, Hilden, Germany) with 20pmol of each primer in a final reaction volume of 50 µl; all primer for kappa plus all primer for the lambda light chain are being used in multiplex in one common assay. PCR cycling conditions are as follows: initial denaturation 15min at 95°C; 50 cycles of 94°C for 30 seconds, 58°C for 90s, and 72°C for 90s, final extension of 10 min.

Agarose Gel Electrophoresis

PCR products were separated on 2% agarose gel (AMRESCO, Solon, Ohio, USA) in 1x TAE buffer, and visualized under UV light by direct

staining with SybrSafe (Invitrogen, Carlsbad, CA, USA) in a 1:20.000 dilution. A 100bp DNA ladder (Promega, Madison, WI, USA) was used as length standard.

NGS Deep Sequencing

Library Preparation and Sequencing

PCR products are purified using standard procedures. Patient-specific indices are ligated to either ends of the purified PCR product (**Error! Reference source not found.**) employing the Illumina TruSeq library preparation protocol (Illumina Corp., San Diego, CA, USA) according to the instructions of the manufacturer, allowing parallel NGS analysis of multiple patients which leads to a significant reduction of the costs.

The quality of each library is tested on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) prior to Polony-PCR (Polymerase-generated-colony) and subsequent sequencing on an Illumina MiSeq (Illumina Corp., San Diego, CA, USA) next-generation sequencing system using the 2x250bp protocol.

Data Analysis

Read 1 and read 2 sequences are overlapped. For this, read 2 is reversely complemented and aligned against read 1. Only alignments in which

85% of the bases in the overlap are similar to each other are further processed. Bases that do not match between read 1 and read 2 are further inspected. If the quality score of one base exceeds 25 whereas the other quality score is below 20, then the base with higher q-value is accepted. Else, the corresponding base of the alignment is changed to “N”. Sequences exceeding 5 Ns are excluded from further analysis.

Using the proprietary “BWA” tool with default parameters, the obtained sequences are mapped against the genomic sequences of all human VH, kappa, and lambda alleles, respectively, obtained from current databases (BLASTn, V-base directory). Reads with more than 100 bp mapped are retained.

Mapped sequences are organized into clusters based on similarity; one cluster consists initially of the most frequent variant among a number of similar sequences. By tolerating a maximum of two to three bases difference, further sequences are grouped to these clusters.

For evaluation of the method, the clonotypic sequence that had been determined upon initial diagnosis is identified among these clusters based on the best match, revealing the number of corresponding reads in the sample as proportion of the total number of mapped reads. For application in clinical routine, it is obsolete to determine the VDJ ‘fingerprint’ of the tumor clone, instead, the presence of a clonal sequence cluster matching the afore mentioned quality criteria is sufficient.

Results

Specificity and Sensitivity Testing

Specificity of the PCR amplification has been verified by PCR amplification of a highly polyclonal DNA template derived from five healthy donors (Fig. 1), followed by NGS-analysis and subsequent alignment to current databases (BLASTn, V-base directory): all seven *IgH* allele families are being amplified at the expected proportion (Table 4) as well as all relevant *IGKL* and *IGLL* alleles (data not shown).

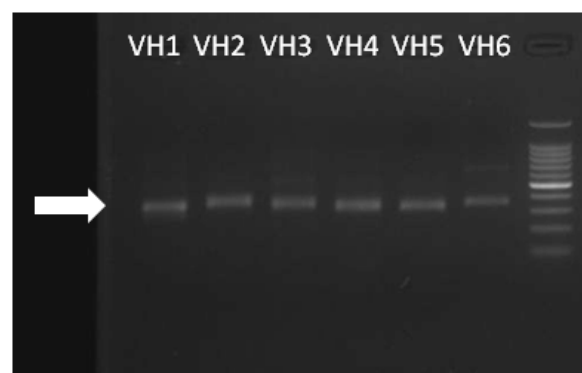


Fig. 1: IgH PCR using family-specific FR1 and consensus JH-primer amplifies all VH allele families from polyclonal DNA; the VH7 family is co-amplified using the VH6 primer

Table 4: Rearranged VH alleles from all seven families can be amplified from 40ng of polyclonal DNA pooled from five donors. The different percentages reflect the proportion of the respective allele among all reads and go well together with the expected representation of the individual VH families in a polyclonal sample.

VH family	Allelic representation [%] of mapped sequences
VH1	20,79
VH2	1,88
VH3	44,54
VH4	23,01
VH5	6,62
VH6	0,64
VH7	2,52

Sensitivity of the PCR reaction has been demonstrated in 5 Myeloma and Burkitt Lymphoma cell lines rearranged for VH1 through VH5 [U266 (ATCC No. TIB-196), BL29 (Rowe, et al., 1987), CA46 (DSMZ-No. ACC-73), and lymphoblastoid cell lines P69 and G57], and 2 Myeloma cell lines rearranged for kappa and lambda light chain, respectively [OPM-2 (IgG Lambda; DSMZ-No. ACC-50), NCI (IgA Kappa; ATCC No. CRL-9068)].

We reach a sensitivity of 6pg clonal DNA, i.e., the PCR protocol is able to detect a single tumor cell (Fig. 2). However, since the total amount of DNA in a PCR reaction is limited to approx. 2µg,

and considering a coverage of at least 300.000 at 7 million reads, the NGS technology is able to detect 1 tumor cell among 330.000 normal B cells within one single PCR reaction (6pg target DNA among 2µg polyclonal background DNA), already exceeding the sensitivity of flow cytometry by two orders of magnitude. Analysis of triplicates would increase the sensitivity to approximately 10^{-6} .

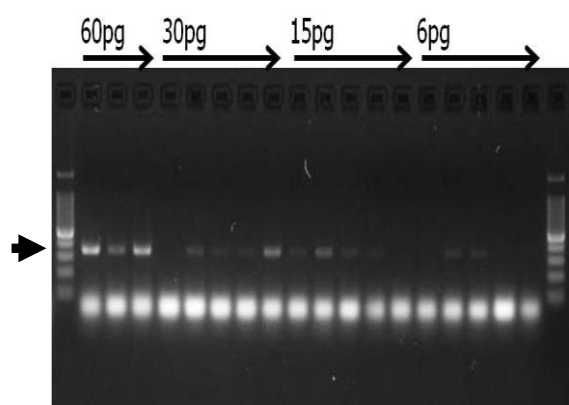


Fig. 2: Sensitivity of the IgH-PCR reaches single cell level, as demonstrated by successful amplification of 6pg DNA of the myeloma cell line G57.

Library Preparation and Cluster Analysis

As a representative example, Fig. 3 shows the quality inspection of three TrueSeq libraries on the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) after

multiplex PCR amplification of 600, 60, and 6pg of clonal DNA from myeloma cell line G57 spiked into 2µg of polyclonal genomic DNA.

A typical example how clonal sequence variants map onto a reference IgH sequence in course of cluster analysis is shown in **Error! Reference source not found.** By this procedure, all sequences are mapped to the most similar reference IgH sequence and aligned to one another according to the number of variations within the sequence, thus revealing any cluster of clonal sequence, which eliminates the necessity to determine the sequence of the primary tumor clone prior to therapy.

Consistency of Amplification Efficiency

For identification of clonotypic sequences based on their number of reads, amplification of the different allele families has to be unbiased. The comparability of the amplification efficiency for each individual VH-allele family has been tested by co-amplification of equal amounts of DNA from myeloma cell line U266 for the VH1 family, LCL P69 for VH2, Burkitt Lymphoma cell lines Raji, BL29, and CA46 for IgH families VH3, VH4 and VH5, respectively. All investigated VH allele families exhibit a comparable amplification efficiency (Table 5) underlining that our approach allows indeed unbiased amplification of the different VH families.

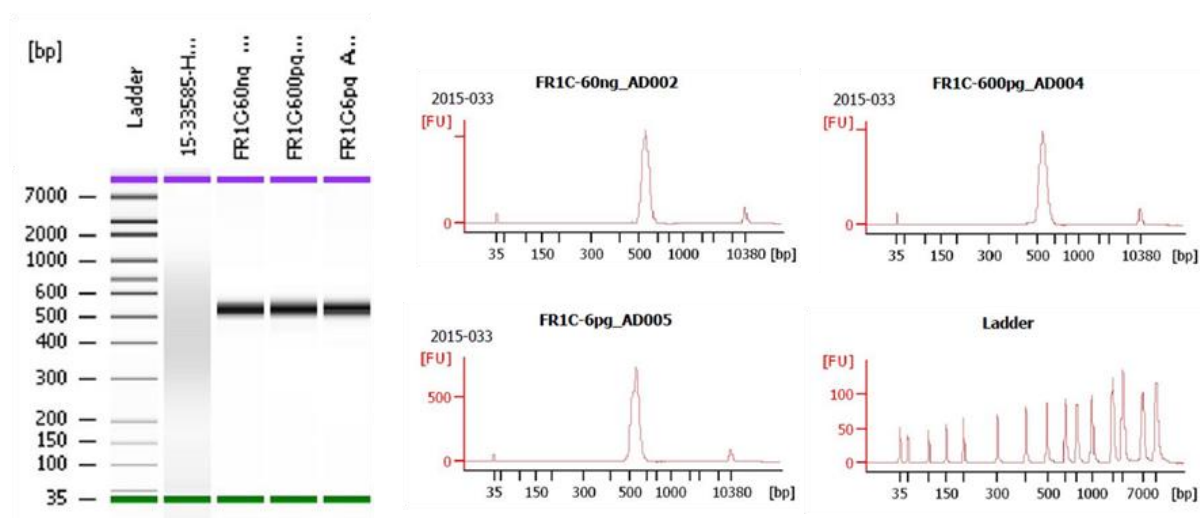


Fig. 3: Quality inspection of the libraries derived from three different PCR products

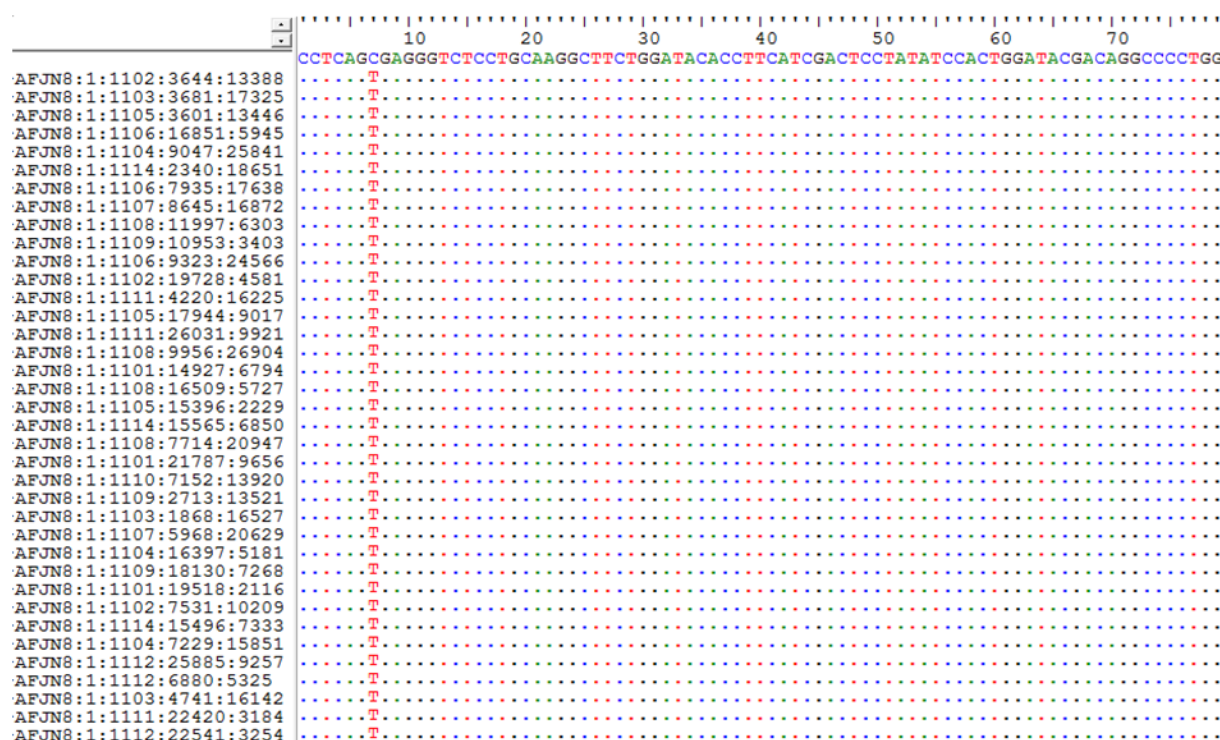
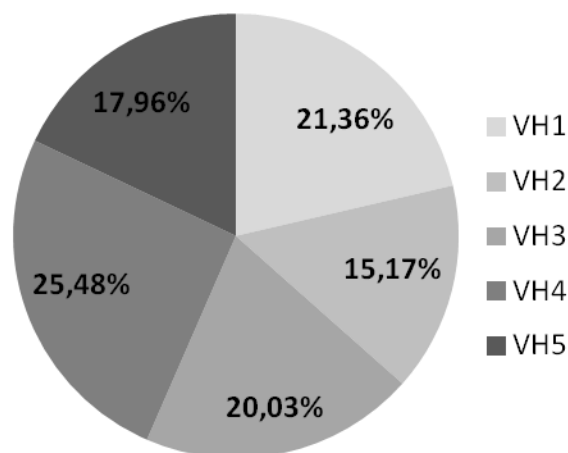


Fig. 4: Representative alignment of mapped sequences to the most frequent sequence variant of the cluster

Table 5: Total and proportional number of mapped reads resulting from multiplex amplification of the IgH locus from equal amounts of clonal DNAs for VH families VH1 through VH5.

VH-family	mapped reads	%
VH1	295.256	21,36%
VH2	209.716	15,17%
VH3	276.801	20,03%
VH4	352.174	25,48%
VH5	248.255	17,96%
total	1.382.202	



Linearity of Amplification

Linear correlation between the amount of template used and the number of clonotypic reads was tested in order to allow for subsequent quantification of the initial template load, i.e., the residual tumor mass. We spiked 600pg, 6ng, 60ng, and 600ng of clonal DNA derived from a myeloma cell line into 1µg of polyclonal DNA obtained from a healthy donor prior to amplification. Indeed, the obtained percentage of

clonotypic reads correlates very well with the initial template amount, allowing for quantification of residual tumor cells by order of magnitude. Each assay yielded more than 30.000 mapped reads, of which 4 were clonotypic for the 600pg assay, 71 for 6ng, 664 for 60ng, and 2657 for 600ng clonal DNA (Table 6). Thus, as the amount of clonal template increases by factor ten, the clonotypic reads as well go up by one order of magnitude.

Within the investigated range of template we achieved a correlation coefficient of $r = 0.9881141$. Hence, our approach as presented here is suitable to reliably detect and quantify any clonal cell population of B-cell origin in the peripheral blood. Quantification of clonal DNA can be enhanced using a reference DNA of known sequence as internal control in each assay.

Table 6: Clonotypic read numbers of serial dilutions of template show excellent correlation to the amount of template used ($r = 0.9881141$)

template	mapped reads	clonotypic reads	proportion
0,6ng	32.014	4	0,0125%
6ng	32.022	71	0,2217%
60ng	31.268	664	2,1236%
600ng	31.927	2.657	8,3221%

Discussion

In this study, we demonstrate that clonal populations of residual cells of B-cell origin, for instance circulating MMPCs in MM patients can successfully be detected and quantified by an unbiased amplification of the entire repertoire of rearranged *IgH* genes, followed by NGS-based massive-parallel ultra-deep sequencing and cluster analysis. Allele drop outs of patients with non-amplifiable *IgH* sequences are avoided by using a two-tiered approach of *IgH* quantification in the first line and multiplexed kappa and lambda light chain quantification in the second line.

This approach eliminates the need to design tumor-clone specific primer and to establish the respective PCR protocol for each patient, rendering NGS-based MRD screening in MM eligible for any clinical molecular diagnostic laboratory, in contrast to qASO-PCR. Moreover, it also eliminates the cumbersome limiting dilution assay required by qASO-PCR, in which multiple replicates are tested at several different dilutions and the proportion of positive and negative results at each dilution level is used to calculate the likely frequency of clonal cells in the starting sample.

Amplification of any rearranged *IgH* sequence provides another stark advantage over qASO-

PCR: cluster analysis, i.e., arrangement of mapped sequences by their similarity to one another and to reference sequences, does not require knowing the sequence of the initial tumor clone, eliminating the need for characterizing the clonotypic VDJ-rearrangement by Sanger sequencing prior to therapy. In contrast to qASO-PCR our approach is also apt to detect any significant clonal population even if they underwent a second round of rearrangement, as well as secondarily occurring tumor clones. This aspect is also critic for MRD screening in B-CLL, since ongoing somatic hypermutations may lead to false negative results when employing qASO-PCR. It goes without saying, that the versatile technique presented here allows MRD screening of any B-cell neoplasia.

We could demonstrate that the primer sets that we designed allow unbiased amplification of all VH families. The good linearity of the unbiased amplification allows reliable quantification of the residual tumor mass within an order of magnitude. With a coverage of at least 300.000 at 7 million reads our approach allows detection of one tumor cell among 330.000 normal B cells (10^{-5}) out of a single PCR reaction (6pg target DNA spiked into 2 μ g polyclonal background DNA). While NGS platforms have the sensitivity to detect 1 tumor cell among a background of 1 million polyclonal cells, the method is limited by the possible input amount of DNA in a PCR reaction. Anyhow, analysis of triplicates would further increase the sensitivity to approximately 10^{-6} .

Martinez-Lopez et al. (17) also achieved a sensitivity of 10^{-6} in the same way while identifying clonal rearrangements as sequences that occurred at a high-frequency of at least 5% in the baseline BM aspirate.

Yoa et al. (18), using the commercial LymphoTrack *IgH* and *IgK* assays (Invivoscribe Inc., San Diego, CA, USA), also showed a sensitivity of 10^{-5} at 1 million sequencing reads using the LymphoTrack-MiSeq platform by spiking 0.001% tumor DNA from plasma cells in 1 μ g polyclonal DNA. However, they had to define the MRD level as the mean MRD burden of triplicate measurements, indicating variations between individual quantifications, which would be a direct consequence of non-homogenous amplification efficiencies.

Another advantage of our approach versus the LymphoTrack system is that in case of allele-dropout of the IgH PCR, our method allows detection of both rearranged kappa and lambda light chains in a single multiplex PCR reaction, while Invivoscribe Inc. only offers an IgK assay, escaping lambda rearranged tumor clones.

Based on these data, we'd like to suggest the following workflow (Fig.5) for MRD screening in MM using targeted massive-parallel ultra-deep-sequencing by NGS: peripheral EDTA anti-coagulated peripheral blood is collected from patients and subjected to DNA extraction using a standard method, followed by amplification of the entire repertoire of rearranged immunoglobulin heavy-chain genes in the sample, representing all circulating normal B-lymphocytes and plasma cells including putative tumor cells. Successful amplification is verified by agarose gel electrophoresis.

After purifying the PCR product by a standard method, patient-specific indices are ligated to the PCR product using the Illumina TruSeq DNA library preparation protocol according to the instructions of the manufacturer (Illumina Corp., San Diego, CA, USA). The quality of each library has to be tested, for instance on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

Data obtained by ultra-deep, massive-parallel sequencing, for instance on the Illumina MiSeq next-generation sequencing platform (Illumina Corp., San Diego, CA, USA) are mapped onto reference immunoglobulin sequences obtained from public databases and clustered in order to detect clonal sequences. The percentage of clonal sequences from the total number of mapped reads in proportion to the total amount of DNA used for PCR amplification allows determining the residual tumor cell mass within an order of magnitude.

A polyclonal result, i.e., cluster analysis fails to reveal a prominent clonal sequence population, indicates no detectable MRD. However, this may only be achieved in patients after allogeneic BMT, since almost all patients following autologous stem cell transplantation remain persistently PCR-positive. Hence, such a result for patients after autologous BMT or patients who didn't receive BMT at all, hints to a so-called allele-dropout, i.e., an amplification-

refractory IgH allele of the tumor clone, which is a rare but known complication for MRD in B-cell malignancies. Therefore, in such a case, the whole assay is repeated starting from amplification of the kappa and lambda light-chains, respectively.

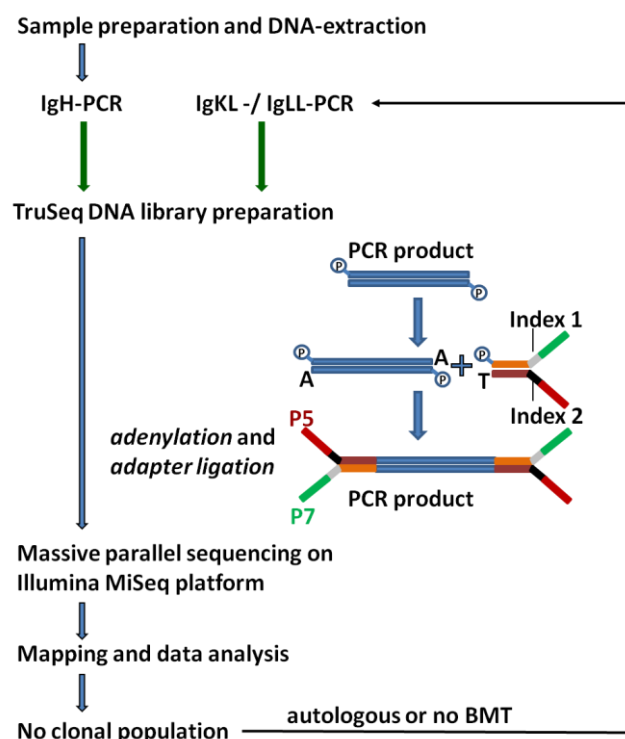


Fig. 5: Suggested workflow for MRD screening in MM by NGS deep sequencing

In summary, this approach eliminates the three main drawbacks of qASO-PCR, since characterization of the genetic fingerprint of the tumor is not required, tumor clones undergoing further mutations and secondary tumors will not escape detection, and the procedure is eligible to any molecular diagnostic laboratory. Apart from this, NGS-based analysis of a standardized PCR reaction allows high throughput analysis at a short turnaround time.

Finally, we should consider the costs of NGS as compared to qASO-PCR. The low costs of the limited-dilution assay underestimate the true costs, which have to include as well Sanger-sequencing of the clonotypic VDJ rearrangement, reagents and labor of establishing the patient-specific PCR protocol which in particular includes the rather high salary for a dedicated highly qualified and experienced researcher. The NGS-based method in contrast

can be implemented by any qualified laboratory staff and pooling multiple patients on one sequencing chip significantly reduces the costs.

Acknowledgments

The authors would like to thank the laboratory of Prof. Dr. M. Pfreundschuh at the José Carreras Center, University Hospital Homburg/Saar, Germany and Prof. Dr. Hartmut Goldschmidt and Dr. Stefanie Huhn from the Clinic of Int. Med V, University Hospital Heidelberg, Germany for supplying cell lines as well as for their valuable advice and critical review of the project.

Disclaimer

The work was sponsored by and carried out at the Institute for Immunology and Genetics, Kaiserslautern, Germany, whereto the first author is affiliated as freelancer, in collaboration with the Clinic of Int. Med V, University Hospital Heidelberg, Germany, during a sabbatical year.

References

1. van Nieuwenhuijzen N, Spaan I, Raymakers R, Peperzak V. From MGUS to multiple myeloma, a paradigm for clonal evolution of premalignant cells. *Cancer Res.* 2018; p. 2449–2456.
2. Nowakowski GS, Witzig TE, Dingli D, Tracz MJ, Gertz MA, Lacy MQ, et al. Circulating plasma cells detected by flow cytometry as a predictor of survival in 302 patients with newly diagnosed multiple myeloma. *Blood.* 2005; p. 2276-2279.
3. Kyle R, Gertz M, Witzig T. Review of 1027 patients with newly diagnosed multiple myeloma. *Mayo Clin Proc.* 2003; p. 21-33.
4. Witzig T, Gertz M, Pineda A, Kyle R, Greipp P. Detection of monoclonal plasma cells in the peripheral blood stem cell harvests of patients with multiple myeloma. *Br J Haematol.* 1995; p. 640–642.
5. Bird J, Bloxham D, Samson D, Marcus R, Russell N, Kelsey S, et al. Molecular detection of clonally rearranged cells in peripheral blood progenitor cell harvests from multiple myeloma patients. *Br J Haematol.* 1994; p. 110-6.
6. Craig J, Langlands K, Parker A, Anthony R. Molecular detection of tumor contamination in peripheral blood stem cell harvests. *Exp Hematol.* 1994; p. 898-902.
7. Mariette X, Feraud J, Brouet J. Myeloma cell contamination of peripheral blood stem cell grafts in patients with multiple myeloma treated by high-dose therapy. *Bone Marrow Transplant.* 1994; p. 47-50.
8. Dreyfus F, Ribrag V, Leblond V, Ravaud P, Melle J, Quarre M, et al. Detection of malignant B cells in peripheral blood stem cell collections after chemotherapy in patients with multiple myeloma. *Bone Marrow Transplant.* 1995; p. 707-11.
9. Gertz M, TWitzig T, Pineda A, Greipp P, Kyle R, Litzow M. Monoclonal plasma cells in the blood stem cell harvest from patients with multiple myeloma are associated with shortened relapse-free survival after transplantation. *Bone Marrow Transplantation.* 1997; p. 337–342.
10. Kiel K, Cremer F, Rottenburger C, Kallmeyer C, Ehrbrecht E, Atzberger A, et al. Analysis of circulating tumor cells in patients with multiple myeloma during the course of high-dose therapy with peripheral blood stem cell transplantation. *Bone Marrow Transplantation.* 1999; p. 1019–1027.
11. Cremer F, Ehrbrecht E, Kiel K, Benner A, Hegenbart U, Ho A, et al. Evaluation of the kinetics of the bone marrow tumor load in the course of sequential high-dose therapy assessed by quantitative PCR as a predictive parameter in patients with multiple myeloma. *Bone Marrow Transplantation.* 2000; p. 851–858.
12. Bakkus MH, Bouko Y, Samson D, Apperley JF, Thielemans K, Van Camp B, et al. Post-transplantation tumour load in bone marrow, as assessed by quantitative ASO-PCR, is a prognostic parameter in multiple myeloma. *British Journal of Haematology.* 2004; p. 665–674.
13. Rawstron AC, Child CA, de Tute RM, Davies FE, Gregory WM, Bell SE, et al. Minimal Residual Disease Assessed by Multiparameter Flow Cytometry in

- Multiple Myeloma: Impact on Outcome in the Medical Research Council Myeloma IX Study. *J Clin Oncol*. 2013: p. 2540-2547.
14. Dingli D, Nowakowski GS, Dispenzieri A, Lacy MQ, Hayman SR, Rajkumar SV, et al. Flow cytometric detection of circulating myeloma cells before transplantation in patients with multiple myeloma: a simple risk stratification system. *Blood*. 2006: p. 3384-3388.
15. Puig N, Sarasquete M, Balanzategui A, Martinez J, Paiva B, Garcia H, et al. Critical evaluation of ASO RQ-PCR for minimal residual disease evaluation in multiple myeloma. A comparative analysis with flow cytometry. *Leukemia*. 2013: p. 1-7.
16. García-Sanz R, López-Pérez R, Langerak A, González D, Chillón M, Balanzategui A. Heteroduplex PCR analysis of rearranged immunoglobulin genes for clonality assessment in multiple myeloma. *Haematologica*. 1999: p. 328-335.
17. Martinez-Lopez J, Lahuerta JJ, Pepin F, González M, Barrio S, Ayala R. Prognostic value of deep sequencing method for minimal residual disease detection in multiple myeloma. *Blood*. 2014: p. 3073-3079.
18. Yao Q, Bai Y, Orfao A, Chim CS. Standardized Minimal Residual Disease Detection by Next-Generation Sequencing in Multiple Myeloma. *Frontiers in Oncology*. 2019: p. 449.
19. Ha J, Lee H, Shin S, Cho H, Chung H, Jang JE. Ig Gene Clonality Analysis Using Next-Generation Sequencing for Improved Minimal Residual Disease Detection with Significant Prognostic Value in Multiple Myeloma Patients. *J Mol Diagn*. 2022: p. 48-56.
20. Katagiri S, Yonezawa T, Kuyama J, Kanayama Y, Nishida K, Abe T, et al. Two distinct human myeloma cell lines originating from one patient with myeloma. *Int J Cancer*. 1985: p. 241-6.
21. Nilsson K, Bennich H, Johansson S, Pontén J. Established immunoglobulin producing myeloma (IgE) and lymphoblastoid (IgG) cell lines from an IgE myeloma patient. *Clin Exp Immunol*. 1970: p. 477-89.
22. Gazdar A, Oie H, Kirsch I, Hollis G. Establishment and characterization of a human plasma cell myeloma culture having a rearranged cellular myc proto-oncogene. *Blood*. 1986: p. 1542-49.
23. Pulvertaft J. Cytology of Burkitt's Tumour (African Lymphoma). *Lancet*. 1964: p. 238-40.
24. Rowe M, Rowe DT, Gregory CD, Young LS, Farrell PJ, Rupani H, et al. Differences in cell growth phenotype reflect novel patterns of Epstein-Barr virus latent gene expression in Burkitt's lymphoma cells. *The EMBO Journal*. 1987: p. 2743-51.
25. Magrath I, Pizzo P, Whang-Peng J, Douglass E, Alabaster O, Gerber P, et al. Characterization of lymphoma-derived cell lines: comparison of cell lines positive and negative for Epstein-Barr virus nuclear antigen. I. Physical, cytogenetic, and growth characteristics. *J Natl Cancer Inst*. 1980: p. 465-76.
26. Magrath I, Freeman C, Pizzo P, Gadek J, Jaffe E, Santaella M, et al. Characterization of lymphoma-derived cell lines: comparison of cell lines positive and negative for Epstein-Barr virus nuclear antigen. II. Surface markers. *J Natl Cancer Inst*. 1980: p. 477-83.
27. van Dongen J, Langerak A, Brüggemann M, Evans P, Hummel M, Lavender F, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 2003: p. 2257-2317.