SUPPLEMENTARY INFORMATION S1

Case number analysis

Based on the estimated 0.39% (1/258) prevalence of CMMRD in malignancy-free children who are suspected of NF1 but lack either *NF1* or *SPRED1* pathogenic variants (PVs), a minimum of 415 patients must be screened for CMMRD to reach a probability of detection >0.8 (Figure S1). This prevalence estimation (for details see Suerink *et al*[Supp.Ref.1]) is based on an assumed CMMRD incidence of one per million newborns to unrelated parents, the assumption that 50% of CMMRD patients have >5 CALMs, and a published detection rate of

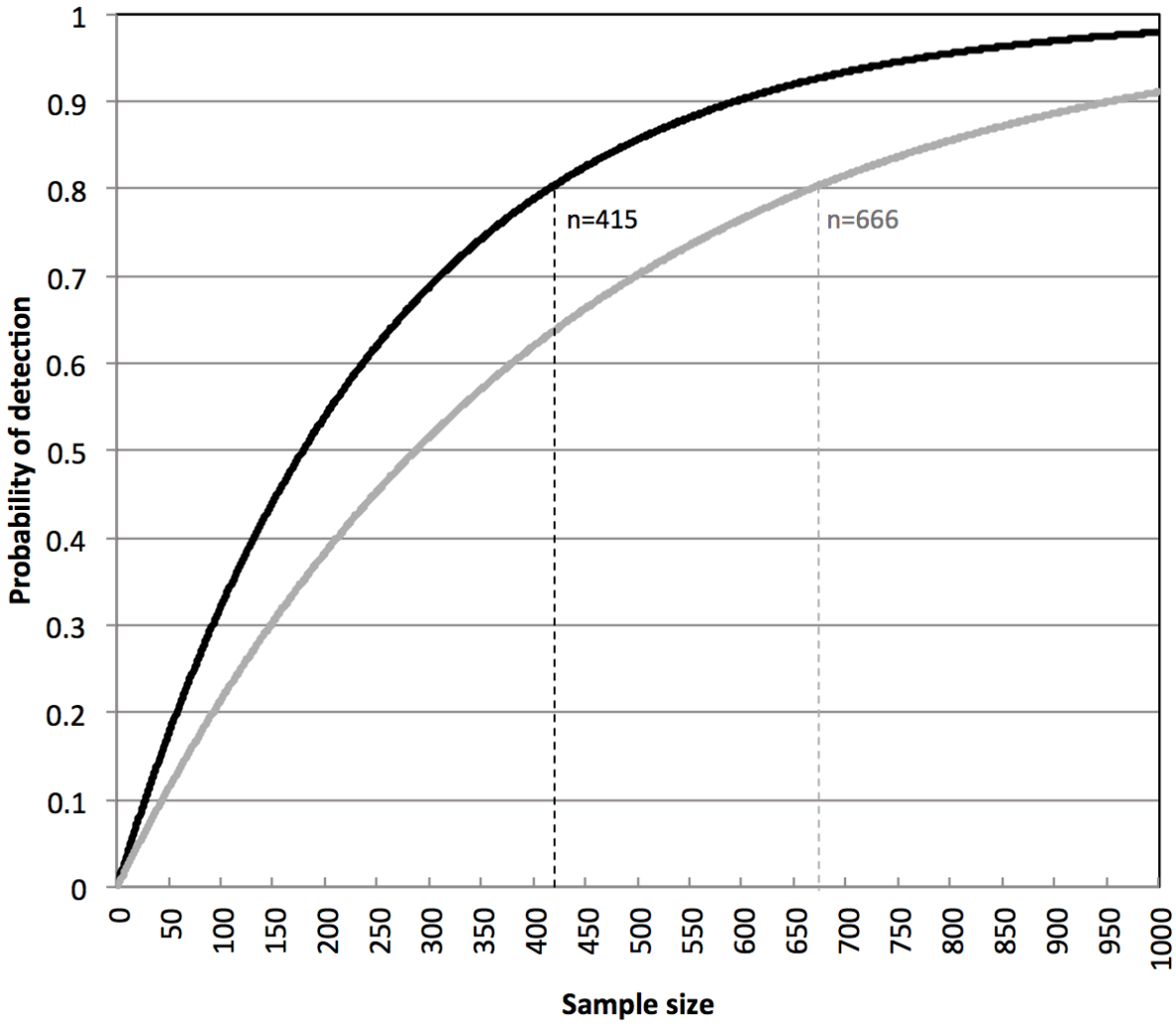

Figure S1: Case number analysis. The traces represent the probability of detection related to the sample size. Black and gray traces show the number of patients to be tested based on a prevalence of 0.39% (1/258) and 0.24% (1/414), respectively. Dashed lines, following the same color scheme, denote the number of samples that should be tested to reach a probability of detection of 0.8.

56.4% for a *NF1* or a *SPRED1* PV in sporadic patients with >5 CALMs irrespective of whether they have additional NIH NF1 features[Supp.Ref.2]. In the same study, sporadic patients with >5 CALMs, with or without freckling, but no additional NIH features had a lower PV rate of 44.6%[Supp.Ref.2]. Taking into account that the majority of CMMRD patients have CALMs with or without freckling but no additional NIH feature, it is more conservative to calculate the CMMRD prevalence using the lower detection rate, which gives a prevalence of 0.24% (1/414). Based on this lower prevalence, the case number analysis shows that 666 patients need to be screened to also reach a probability of detection >0.8 (Figure S1).

SUPPLEMENTARY INFORMATION S2

In this study, we screened for constitutional mismatch repair deficiency (CMMRD) in a cohort of 752 patients using a recently published microsatellite instability (MSI) assay[Supp.Ref.3]. Given the large number of patients, we updated the MSI scoring method and introduced additional quality control (QC) criteria to minimize false positive and negative results.

Assessing microsatellite instability score reliability by single molecule sequence count

The dataset previously generated at Newcastle University in the study of Gallon *et al*[Supp.Ref.3] was used to assess the reliability of MSI score based on the number of single molecule sequences (smSequences) analyzed. To generate the MSI score, each microsatellite marker is assigned a probability by comparing the proportion of its smSequences that have a WT microsatellite length (prWT) to a beta distribution derived from a reference set of 40 controls (for details see Gallon *et al*[Supp.Ref.3]). Here, this probability will be referred to as βWT. βWT was compared to smSequence counts from 106 control samples analyzed in Newcastle. When <100 smSequences were analyzed, markers were more likely to have a low βWT, as can be seen by the decrease in 5^{th}, 25^{th} and 50^{th} percentiles of βWT relative to when >100 smSequences are analyzed (Figure S2.1). Therefore, to ensure that low smSequence counts would not affect sample classification, the method was updated so that markers with <100 smSequences would be excluded from the MSI score calculation. Furthermore, to minimize the impact of these marker exclusions and of low smSequence numbers at sample level, MSI scores would fail to pass QC if i) >2 markers were excluded due to having fewer than 100 smSequences; and, ii) the mean smSequence count was <200.

Single nucleotide polymorphisms can be used to detect sample contamination

Within the dataset generated at Newcastle University by Gallon *et al*[Supp.Ref.3], a control gDNA sample, MLPA101 (used to assess inter-run variation of the MSI assay), had an MSI score >2.00 in one of its repeats (MLPA101R3), despite all other assays of MLPA101 having an MSI score <0.50. Subsequent analysis of MLPA101R3 found that 6 of the SNPs, which are sequenced with the microsatellites, included novel alleles (frequencies <0.12), that were not found

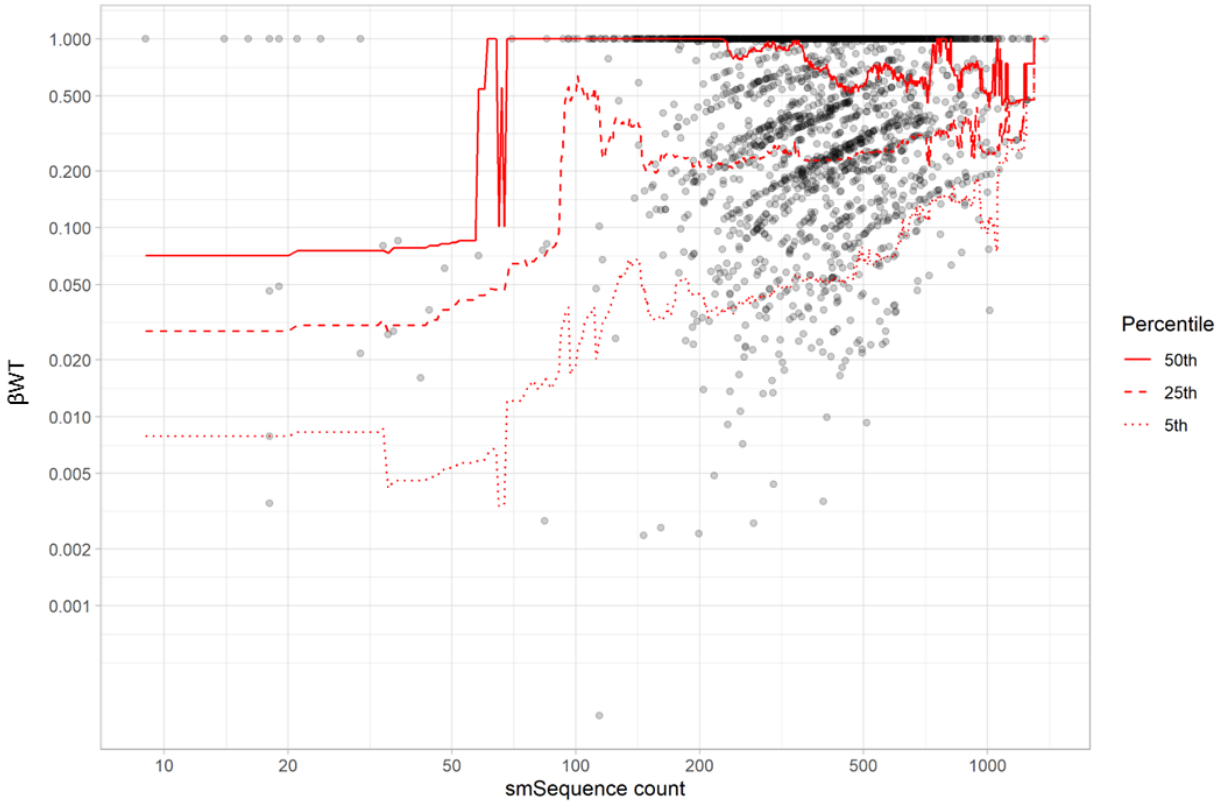

Figure S2.1: βWT compared to smSequence count for 24 microsatellite markers analyzed in 106 control samples. The 50th, 25th and 5th percentiles (represented by the red line overlays) are calculated from the βWT within a range ±50 of the abscissa smSequence count.

in the other MLPA101 analyses. In addition, 5 heterozygous SNPs showed a shift in allele frequency from approximately 0.5 in the other MLPA101 repeats to >0.6 in MLPA101R3. In some markers, these alleles were associated with novel or increased-frequency microsatellite length variants (Figure S2.2), indicating that contamination was influencing MSI score. Therefore, a flag for SNP minor allele frequencies >0.01 and <0.40 was built into the scoring method to highlight markers and samples for visual inspection of smSequence data. An additional QC criterion was introduced such that an MSI score would only pass QC and be considered reliable if there was deemed to be no evidence of contamination on visual inspection of flagged smSequence data.

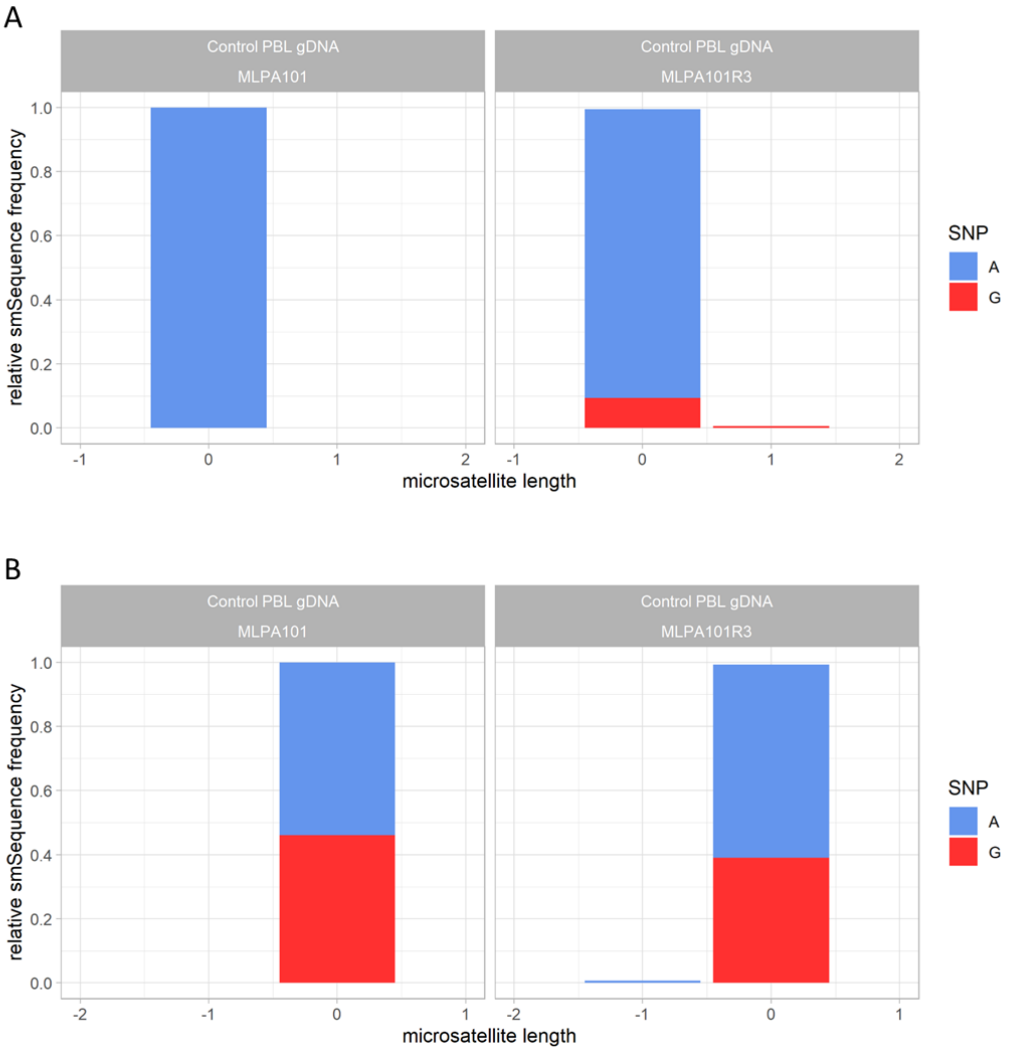

Figure S2.2: SNP allele and microsatellite length distribution in smSequences for **A** marker GM01, and **B** marker GM22, comparing results from the first assay of control sample MLPA101 (MLPA101) with the results from a contaminated repeat of MLPA101 (MLPA101R3).

Port of the microsatellite instability assay

The MSI assay was ported to the Medical University of Innsbruck and initially evaluated using a pilot cohort of 14 CMMRD and 13 control peripheral blood leukocyte (PBL) gDNAs previously analyzed in Newcastle. Samples were amplified, sequenced and scored according to Gallon *et al*[Supp.Ref.3], except that the EnvStats R package was used rather than ExtDist due to incompatibility with R v3.5.3. By the classification threshold of MSI score >2.00, there was 100% concordance in sample classification between the two laboratories. However, one sample, DNA93, had an exceptionally high score difference of 43.53 (Innsbruck score = 10.19, Newcastle score = 53.72). This difference was explained by one marker, LR52, with a germline

length variant that had been excluded from scoring in Innsbruck as a probable germline variant (prWT = 0.50) but had been included in Newcastle (prWT 0.61) as it was above the germline exclusion threshold of prWT ≤0.60. With the exclusion of LR52, DNA93's Newcastle score dropped to 10.91. A more stringent prWT threshold of <0.75 was therefore introduced. This threshold would be highly unlikely to exclude markers with decreased prWT due somatic variation in CMMRD samples as prWT was consistently >0.83 for all non-germline variant markers in the CMMRD samples analyzed in Gallon *et al*[Supp.Ref.3]. To further explore any differences in MSI assay performance between the two laboratories, 58 control PBL gDNAs were analyzed in Innsbruck and scored by the method of Gallon *et al*[Supp.Ref.3]. The Innsbruck control MSI scores were compared to MSI scores from the 40 controls analyzed in Newcastle that make up the reference set, against which both sets of controls were scored. There was no significant difference in scores between the two laboratories, although there was a trend for slightly lower scores in Innsbruck (Figure S2.3, p = 0.09). To account for this and to make the reference set more robust, prWT data from these 58 controls were combined with the original 40 controls into a new reference set of 98 controls.

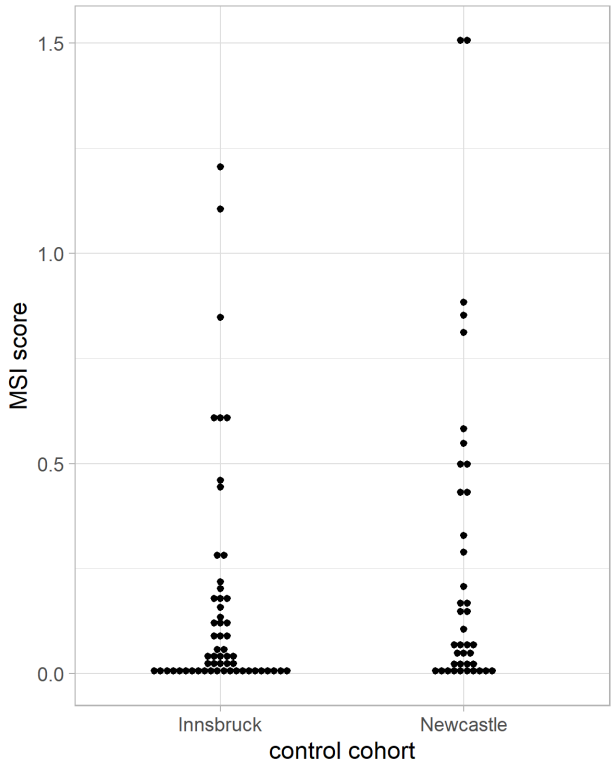

Figure S2.3: Comparison of MSI scores of control samples analysed in Innsbruck and Newcastle.

SUPPLEMENTARY INFORMATION S3

Reagent costs per sample for library preparation

Item	Cost per item (£)	No. of samples	Cost per sample (£)
MIPs (24 MIPs per reaction)	240.00	500000	0.01
T4 polynucleotide kinase (2500 Units, 10U/ul)	170.00	500000	0.01
T4 DNA ligase buffer (6ml)	16.00	500000	0.01
Ampligase DNA ligase (1000 Units, 5U/ul)	142.00	200	0.71
Herculase DNA polymerase 2 (400 reaction kit)	290.00	225	1.29
Exonuclease I (15,000 Units, 20U/ul)	207.00	750	0.28
Exonuclease III (25,000 Units, 100U/ul)	182.00	250	0.73
MIP sample indexing reverse primers (200 oligos)	1700.00	1200000	0.01
MIP forward primer	8.50	3200	0.01
QIAxcel screening kit (2400 samples)[a]	542.00	2400	0.23
AMPure XP DNA cleanup kit (60 ml)	1146.00	1333	0.86
10mM Tris HCl pH8.5 (250ml EB buffer, Qiagen)	32.30	2500	0.01
100% Ethanol (4L, Thermofisher Scientific)	205.42	14285	0.01
Qubit dsDNA HS kit (500 reaction kit)	175.00	500	0.35
Custom sequencing primers (3 oligos)	15.00	18750	0.01
		Total (£)	**4.53**

[a] equivalent cost to analysis by gel electrophoresis

Reagent costs per sample for sequencing

Sequencing kit	Cost per item (£)	No. of samples[a]	Cost per sample (£)
MiSeq v3 Kit (600 cycles, 25 million reads)	1358	174	7.80
MiSeq v2 Kit (300 cycles, 15 million reads)	928	104	8.92
MiSeq v2 Micro Kit (300 cycles, 4 million reads)	390	28	13.93
MiSeq v2 Nano Kit (300 cycles, 1 million reads)	259	7	37.00

[a] using 4000 reads/marker/sample and 24 markers/sample and assuming two thirds of reads are on target

Reagent costs per sample for library preparation and sequencing combined

Sequencing kit	No. of samples[a]	Cost per sample (£)	Cost per sample ($)
MiSeq v3 Kit (600 cycles, 25 million reads)	174	12.33	15.54
MiSeq v2 Kit (300 cycles, 15 million reads)	104	13.45	16.95
MiSeq v2 Micro Kit (300 cycles, 4 million reads)	28	18.45	23.25
MiSeq v2 Nano Kit (300 cycles, 1 million reads)	7	41.53	52.32

[a] using 4000 reads/marker/sample and 24 markers/sample and assuming two thirds of reads are on target

SUPPLEMENTARY REFERENCES

1. Suerink, M. *et al.* Constitutional mismatch repair deficiency as a differential diagnosis of neurofibromatosis type 1: Consensus guidelines for testing a child without malignancy. *Journal of Medical Genetics* **56,** 53–62 (2019).
2. Messiaen, L. *et al.* Clinical and Mutational Spectrum of Neurofibromatosis Type 1–like Syndrome. *JAMA* **302,** 2111–2118 (2009).
3. Gallon, R. *et al.* A sensitive and scalable microsatellite instability assay to diagnose constitutional mismatch repair deficiency by sequencing of peripheral blood leukocytes. *Hum. Mutat.* **40,** 649–655 (2019).