WHITE PAPER

Making Coverage Count: Screening for *CFTR* Mutations in Diverse Populations with Effective Variant Panels

SEPTEMBER 2020

Asuragen®

Making Coverage Count: Screening for *CFTR* Mutations in Diverse Populations with Effective Variant Panels

Overview

Cystic Fibrosis (CF) is among the most prevalent lethal autosomal recessive genetic diseases, with a pan-ethnic frequency of around 1 in 3,500 live births in the United States and Europe[1-4]. While the disease impacts multiple organs, primary morbidity and mortality are associated with progressive loss of lung function and pulmonary disease caused by abnormal thickening of mucous, with an average survival of 40 years in patients with classic CF[1,3,5]. Recently developed treatments have proven effective for many patients[6], contingent on accurate genetic diagnosis. Non-classic CF is less common, with milder effects such as infertility, pancreatitis, and chronic lung and sinus issues[5].

CF is caused by mutations in *CFTR*, a large gene with 27 exons that encodes an important transporter protein. CFTR protein regulates the flow of water in cells that produce mucous, sweat, and other fluids. Pathogenic mutations in both copies of the gene cause a breakdown in this mechanism, causing mucosal thickening responsible for primary symptoms. In the US, nearly 4% (~1/25) of the population are CF carriers[7], with one functional copy of *CFTR* and one copy with one or more pathogenic mutations.

CF carriers occur most frequently in non-Hispanic white (1:25) and Ashkenazi Jewish (1:24) ethnicities, with lower frequencies (1:58 to 1:94) in other ethnicities[8]. For this reason, knowledge of the mutations responsible for CF have historically focused on ethnicities with higher prevalence[3,5]. Modernization of sequencing through NGS has enabled comprehensive assessment of over 2,000 CF mutations with broad ethnic representation, identifying significant heterogeneity between ethnic backgrounds[1,3,5,7]. Recent studies show that typical CF mutation panels recommended by guidelines do not include mutations frequently found in many ethnicities, leading to reduced detection sensitivity in diverse populations[1,7].

In this paper, we will discuss implications of our expanding understanding of CF and its underlying genetics for a broader range of ethnicities, particularly as it pertains to effective testing for all individuals regardless of ethnic background.

Genetic Testing for CF

Genetic testing to identify pathogenic *CFTR* variants in isolated genomic DNA is the standard for carrier testing in adults, and is also used for diagnostic testing both postnatally and in adults[5,8,9]. Because of the large number of possible mutations and size of the *CFTR* gene, testing is often targeted to specific pathogenic variants that occur with high frequency, an approach known as classification-based or targeted testing[4,5,8]. While this enables rapid, cost-effective testing using PCR or similar methods, it requires that targeted variants are representative of the population to be tested in order to maximize coverage[5,8,9], which can be difficult in large, diverse populations such as those found in the US.

An alternative approach is to use NGS sequencing. Sequencing can provide comprehensive testing of *CFTR* variants. However, this approach requires interpretation of the pathogenicity of each identified variant, also known as classification-based reporting, as most documented mutations in the *CFTR* gene are benign or have variable or unknown clinical consequences[5]. Additionally, NGS methods are often more costly and

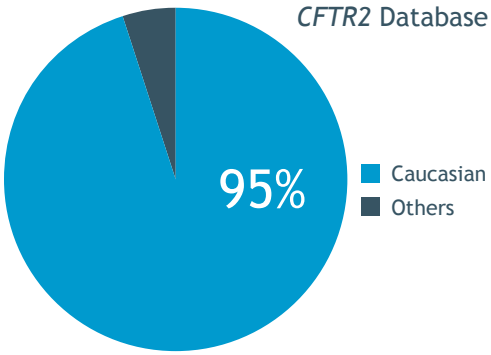

Figure 1A: Ethnic diversity of 89K patients in *CFTR2* database

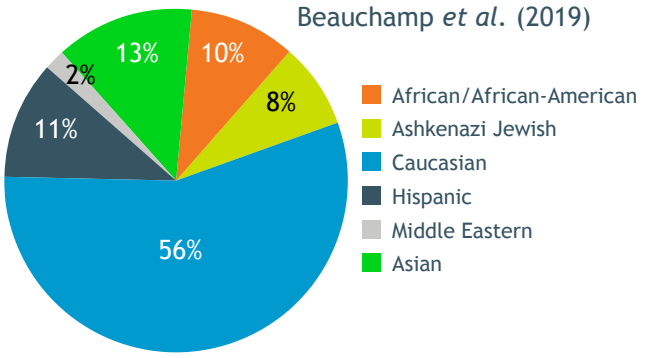

Figure 1B: Ethnic diversity of 115K subjects in US study

Commercial Kit Manufacturer/Panel	Full Product Name	# of Variants Detected	% Coverage, *CFTR2*	% Coverage, US[a]	% Coverage Difference, US[b]
Asuragen	AmplideX PCR/CE *CFTR* Kit	67	92.1%	93.0%	-
Illumina	MiSeqDx Cystic Fibrosis 139 Variant Assay	139	94.3%	87.7%	5.3%
Luminex 71	xTAG Cystic Fibrosis (CFTR) 71 kit v2	71	91.5%	87.0%	6.0%
Agena	iPLEXPro CFTR Panel	74	91.5%	86.9%	6.1%
Elucigene	CF-EU2v1	50	90.5%	86.2%	6.8%
Luminex 60	xTAG Cystic Fibrosis (CFTR) 60 kit v2	60	91.2%	86.1%	6.9%
Luminex 39	xTAG Cystic Fibrosis (CFTR) 39 kit v2	39	89.1%	80.7%	12.3%
CF 23/GenMark	eSensor CF Genotyping Test	23	86.8%	78.9%	14.1%

[a]From Beauchamp *et al*. 2019
[b]% Coverage difference from Asuragen assay based on Beauchamp *et al*. 2019

Figure 1C: Percent coverage of variants detected by commercial kits based on *CFTR2* database and US population frequencies. Sorted highest to lowest by US population coverage.

complex than targeted approaches, and require sophisticated analysis and longer turnaround times. Thus, there is a tradeoff between maximizing *CFTR* coverage using NGS with testing costs, interpretive challenges, skilled labor needs, and turnaround time compared to targeted assays.

In the US, CF carrier screening is recommended for all women considering pregnancy regardless of ethnicity by both American College of Obstetricians and Gynecologists (ACOG) and American the College of Medical Genetics and Genomics (ACMG)[5,8]. As the first genetic disease to have well-established carrier screening guidelines, a core panel of 23 CF mutations (ACMG 23, or CF 23) was established in 2004 and recently reaffirmed as the minimum panel for all targeted CF testing regardless of indication[5,8]. However, detection rates using this panel vary significantly depending on ethnicity, with sensitivity for detection of carriers ranging from nearly 95% to as low as 49%[1,5,8].

Given the high degree of pan-ethnic heterogeneity observed among more than 2,000 known CF mutations, current guidelines suggest including additional mutations in test panels based on the diversity of the intended test population to improve coverage[5,8]. European CF carrier screening guidelines are more direct, indicating that mutations with a frequency of more than 0.5 to 1% in the intended test population should be included in panels[4].

Whereas some commercial CF kits focus solely on the CF 23 panel (i.e. GenMark), many have expanded panels for improved coverage and sensitivity, covering anywhere from 43 to 139 CF variants (Luminex, Illumina, Elucigene, Agena). These kits have largely been designed based on databases like CFTR2, which is one of the largest collections of *CFTR* variants and drawn from >89,000 CF patients. While this database

includes some patients from all ethnicities[5], 95% are non-Hispanic white (**Figure 1A**)[1]. Thus, while most commercial kits have coverage of 90% or more based on variant frequencies according to CFTR2 (**Figure 1C**), these frequencies are biased toward non-Hispanic white ethnicities.

In order to more accurately characterize *CFTR* variant frequencies and incorporate representative population diversity, a recent CF carrier screening study using comprehensive NGS testing examined over 115,000 individuals from a pan-ethnic US population (**Figure 1B**), identifying nearly 4,000 carriers[1]. They found that the CF 23 panel would fail to detect 31% of at-risk individuals, with disproportionate impact on Hispanic, African, and East Asian ethnicities. Based on variant frequencies from this diverse population, coverage of other commercially available panels is similar, ranging from 79-88% (**Figure 1C**). These data suggest that most commercial panels have reduced coverage in diverse populations that disproportionately excludes non-white ethnicities. Based on this finding, the authors conclude that sensitivity (i.e. coverage) and specificity, rather than number of variants, is the most critical variable for evaluating test efficacy[1].

The variants targeted by the AmplideX PCR/CE *CFTR* Kit* (Asuragen) is based on the most current data describing *CFTR* variant representation across different ethnic groups, and the US demographic as a whole. The assay detects 67 pathogenic *CFTR* mutations in two PCR reactions, including SNPs, INDELs, and large exon deletions. The kit also sizes poly-T/TG repeats associated with R117H, which can help resolve non-classic CF or other less common phenotypes. The allele-specific amplicons are generated from purified genomic DNA, resolved using capillary electrophoresis (CE), and analyzed using automated software included with the kit. The workflow, which is similar to that of Asuragen's AmplideX *FMR1*[†,‡] and *SMN1/2* Plus[†] PCR/CE kits, requires less than five hours from sample-to-answer with less than an hour of hands-on time.

While other targeted *CFTR* kits have similar numbers of variants, the Asuragen kit has the highest percent coverage measured against CFTR2 for non-NGS workflows (**Figure 1C**). Designed with diversity in mind, it improves the absolute coverage of CF carriers by 5 to 14% over other targeted *CFTR* assays when examining variant frequencies from a large, ethnically representative population (**Figure 1C**). Thus, the AmplideX PCR/CE *CFTR* Kit* combines the speed, simplicity, and affordability of targeted *CFTR* testing with unmatched coverage guided by the most updated variant prevalence data from different populations. In the next section, we will closely examine CF variants identified in several recent large scale carrier screening studies, and how they impact coverage across targeted *CFTR* panels.

THE MODERN LANDSCAPE OF CF VARIANT DIVERSITY

In recent years, the widespread implementation of NGS for expanded carrier screening has enabled examination of genes involved in diseases like CF with unprecedented depth. These carrier studies are particularly valuable, as the screening population more closely resembles random sampling compared to studies of individuals diagnosed with CF, which are biased by varying disease prevalence among different ethnicities. The results from these screens offer a more accurate reflection of variant diversity in the overall population and within each ethnicity. In two recent studies, nearly 500,000 individuals spanning diverse ethnicities were comprehensively screened for *CFTR* mutations[1,7]. Here, we will examine key *CFTR* variants and associated frequencies identified from these studies and discuss the impact of these findings on commonly used CF panels.

In the Beauchamp *et al*. study, 115,000 individuals in the US were screened for *CFTR* mutations. The ethnic composition of this group closely reflected the US demographic rather than typical CF database frequencies[1], illustrated by **Figure 1B** compared to **Figure 1A**. The results show that the coverage of the most frequent variants within the intended test population is more important than the number of variants in a panel, as several variants with high frequencies in this pan-ethnic population are missed by current panels, with coverage ranging from approximately 79-88% for most commercial offerings (**Figure 1C**).

Table 1: Top 25 *CFTR* variants from carrier screening 115K individuals[a]. Circles indicate variants detected, squares indicate variants detected only with the Asuragen assay. Variants sorted by allele count. See Figure 1C for full product names of the listed commercial kits.

Top variants across all ethnicities[a]			Variant included in commercial kit							
Variants	Allele Count	Allele Frequency (N=3,965)	CF 23/ GenMark	Luminex 39	Luminex 60	Luminex 71	Elucigene	Agena	Illumina	Asuragen
c.1521_1523delCTT (aka F508del)	2096	52.9%	●	●	●	●	●	●	●	●
c.350G>A (R117H)	347	8.8%	●	●	●	●	●	●	●	●
c.3846G>A (W1282*)	152	3.8%	●	●	●	●	●	●	●	●
c.3454G>C (D1152H)	114	2.9%	-	-	●	●	●	●	-	●
c.1624G>T (G542*)	99	2.5%	●	●	●	●	●	●	●	●
c.1652G>A (G551D)	74	1.9%	●	●	●	●	●	●	●	●
c.3209G>A (R1070Q)	71	1.8%	-	-	-	-	-	-	-	■
c.3909C>G (N1303K)	48	1.2%	●	●	●	●	●	●	●	●
c.3718-2477C>T (aka 3849+10kbC>T)	46	1.2%	●	●	●	●	●	●	●	●
c.617T>G (L206W)	36	0.9%	-	-	●	●	●	●	●	●
c.489+1G>T (aka 621+1G>T)	31	0.8%	●	●	●	●	●	●	●	●
c.2657+5G>A (aka 2789+5G>A)	31	0.8%	●	●	●	●	●	●	●	●
c.1367T>C (V456A)	31	0.8%	-	-	-	-	-	-	-	■
c.1657C>T (R553*)	30	0.8%	●	●	●	●	●	●	●	●
c.2988+1G>A (aka 3120+1G>A)	30	0.8%	●	●	●	●	●	●	●	●
c.349C>T (R117C)	27	0.7%	-	-	-	●	●	●	●	●
c.1585-1G>A (aka 1717-1G>A)	26	0.7%	●	●	●	●	●	●	●	●
c.1519_1521delATC (aka I507del)	19	0.5%	●	●	●	●	●	●	●	●
c.2657+2_2657+3insA	19	0.5%	-	-	-	-	-	-	-	■
c.1000C>T (R334W)	16	0.4%	●	●	●	●	●	●	●	●
c.14C>T (P5L)	15	0.4%	-	-	-	-	-	-	-	■
c.254G>A (G85E)	14	0.4%	●	●	●	●	●	●	●	●
c.1040G>C (R347P)	14	0.4%	●	●	●	●	●	●	●	●
c.2052dupA (aka 2184insA)	14	0.4%	-	-	-	-	-	-	●	●
c.1853T>C (I618T)	13	0.3%	-	-	-	-	-	-	-	■

[a]From Beauchamp *et al*. 2019

Dissecting the top 25 pathogenic or likely pathogenic variants from this study across all ethnicities, five variants are covered only with the Asuragen assay (**Table 1**, squares), including one of the top ten most frequently observed variants in the study. While these five variants account for nearly 4% of the overall pan-ethnic allele frequency in the study, they represent less than 0.1% of overall allele frequency from

Table 2: Top 5 *CFTR* variants for each ethnicity from carrier screening 381K individuals[a]. Circles indicate variants detected, squares indicate variants detected only with the Asuragen assay. Variants sorted by allele count. See Figure 1C for full product names of the listed commercial kits.

Variants in top 5 for ≥1 ethnicity[a]		Reported allele frequency for top 5 variants in each ethnicity[a, c]						Variant included in commercial kit							
Variants[b]	Allele Count	Caucasian (N=7,311)	Hispanic (N=1,880)	African American (N=1,271)	East Asian (N=134)	South East Asian (N=187)	Ashkenazi Jewish (N=255)	CF 23/ GenMark	Luminex 39	Luminex 60	Luminex 71	Elucigene	Agena	Illumina	Asuragen
c.1521_1523delCTT (p.F508del)	4812	52.87%	29.84%	23.29%	7.46%	12.83%	22.96%	●	●	●	●	●	●	●	●
c.350G>A (p.R117H)	621	7.62%	3.4%	-	-	-	-	●	●	●	●	●	●	●	●
c.3154T>G (p.F1052V)	344	3.09%	5.32%	-	-	-	7.06%	-	-	-	-	-	-	-	-
c.3454G>C (p.D1152H)	304	2.07%	6.33%	-	-	-	13.33%	-	-	●	●	●	●	-	●
c.1624G>T (p.G542*)	215	1.9%	3.62%	-	-	-	3.14%	●	●	●	●	●	●	●	●
c.1865G>A (p.G622D)	152	-	-	6.77%	14.93%	24.6%	-	-	-	-	●	-	-	-	●
c.2988+1G>A	121	-	-	9.52%	-	-	-	●	●	●	●	●	●	●	●
c.3846G>A (p.W1282*)	74	-	-	-	-	-	29.02%	●	●	●	●	●	●	●	●
c.3209G>A (p.R1070Q)	51	-	-	-	13.43%	17.65%	-	-	-	-	-	-	-	-	■
c.1853T>C (p.I618T)	42	-	-	3.3%	-	-	-	-	-	-	-	-	-	-	■
c.3297C>A (p.F1099L)	32	-	-	2.52%	-	-	-	-	-	-	-	-	-	-	-
c.1558G>A (p.V520I)	15	-	-	-	-	8.02%	-	-	●[d]	●[d]	●[d]	●[d]	●[d]	●[d]	●[d]
c.3205G>A (p.G1069R)	15	-	-	-	11.19%	-	-	-	-	-	-	-	-	-	-
c.1367T>C (p.V456A)	14	-	-	-	-	7.49%	-	-	-	-	-	-	-	-	■
c.2909G>A (p.G970D)	7	-	-	-	5.22%	-	-	-	-	-	-	-	-	-	■
Total % covered[c]:		**67.55**	**48.51**	**45.4**	**52.23**	**70.59**	**75.51**								

[a]From Westemeyer *et al.* 2020
[b]The variants noted here reflect classification as likely pathogenic or pathogenic at the time of manuscript submission. Classifications may change over time.
[c]Only includes results for top 5 most frequent variants in each ethnic group
[d]Detects 1558G>T(V520F)

the CFTR2 database. This frequency difference of more than 10-fold in the large carrier screening study relative to the patient database emphasizes the importance of characterizing variant frequencies in a population that is representative of the intended test population.

Another recent study by Westemeyer *et al.* examined over 380,000 individuals tested over three years for expanded carrier screening[6]. Of these, 98.4% (374,911) were screened for *CFTR*, with sequencing methods used to identify 14,229 CF carriers. The authors found that the CF 23 panel would fail to detect approximately 44% of carriers, even more than the 31% of carriers missed in the Beauchamp study.

Utilizing the diverse screening population and large study size, Westemeyer examined the top five most frequent *CFTR* variants within each ethnic group to further illustrate heterogeneity, as shown in **Table 2**. While top variants were similar between Caucasian, Hispanic, and Ashkenazi Jewish individuals, the most

frequent variants in African American, East Asian, and South East Asian individuals were highly divergent, with approximately 15% of carriers in each ethnic group having a variant not frequently found in other ethnicities.

Of the top variants from each ethnicity in the Westemeyer study, only five of the fifteen variants are included in the CF 23 panel, with most other panels covering six or seven (**Table 2**). In contrast, eleven of these fifteen variants are included in the Asuragen panel, including four variants that are unique to this panel (**Table 2**, squares). Three variants (F1052V, F1099L, and G1069R) are not included in any assay. However, these are of varying clinical significance according to the CFTR2 database, and therefore may not be appropriate for reporting based on current guidelines[5].

Of the four mutations from the Westemeyer study that are unique to the Asuragen panel, three were among the top 25 variants identified in the Beauchamp study. The high frequency of these variants across nearly 500,000 patients tested in both of these studies confirms the utility of including these variants in screening and diagnostic panels for improved detection in diverse populations. While the combined allele frequency of the four variants from Westemeyer is less than 0.05% in CFTR2, they represent 3.3% of African American carriers, 18.7% of East Asian carriers, and 25% of South East Asian carriers identified. This contrast highlights the implicit ethnic bias in evaluating *CFTR* variant frequencies based on the disease population, which is significantly impacted by disease prevalence within each ethnicity.

CLINICAL IMPACT OF *CFTR* HETEROGENEITY

The diversity of *CFTR* variants among ethnic groups identified in these and other studies have significant real-world implications. For carrier testing, reduced variant coverage translates to reduced detection of couples at risk of a child affected by CF (at risk couples, ARC). Detection of ARCs requires identifying pathogenic variants in both parents, compounding differences in percent coverage between panels. The Asuragen assay design has the potential to improve the detection rate of ARCs from 9.6% to 24.2% compared to other panels (**Table 3**). Given that 89% of at-risk couples take action to reduce the risk of CF-affected pregnancy by considering alternatives like *in vitro* fertilization, prenatal diagnostic testing, or adoption[1], this could lead to increased risk of children diagnosed with CF among impacted ethnicities such as African Americans and Asians.

Table 3: Coverage and estimated at risk couple (ARC) detection rate of commercial *CFTR* kits. Sorted highest to lowest by detection rate. See Figure 1C for full product names of the listed commercial kits.

Commercial Kit	% Coverage[a]	At Risk Couple Detection Rate[b]
Asuragen	93.0%	86.5%
Illumina	87.7%	76.9%
Luminex 71	87.0%	75.7%
Agena	86.9%	75.5%
Elucigene	86.2%	74.3%
Luminex 60	86.1%	74.1%
Luminex 39	80.7%	65.1%
CF 23/GenMark	78.9%	62.3%

[a]From Beauchamp et al. 2019
[b]% Coverage squared for carrier alleles in both parents. Predicted at-risk couples (ARC) calculated from *CFTR* mutation coverage assuming fully pathogenic alleles.

Additionally, because symptomatic identification typically delays diagnosis to 14.5 months of age, genetic testing can also significantly expedite diagnosis by detecting CF in asymptomatic patients[10,11]. This is important because timing is critical, as patients that are not diagnosed until symptoms occur have significantly worse clinical outcomes, with higher rates of complications or hospitalizations compared to patients diagnosed by pre-symptomatic newborn screening[10]. Since genetic testing is often used in confirmatory steps of newborn CF screening protocols, the exclusion of variants specific to certain ethnicities on test panels has likely contributed to CF diagnosis at a later age among several non-white ethnicities as compared to whites, illustrating how reduced coverage translates to worse clinical outcomes in minority groups[3,12,13]. Given the recent and ongoing development of targeted treatments for CF that continue to improve patient outcomes when combined with early detection[1,5], the timing of diagnosis is more critical than ever.

CONCLUSION

The evidence provided herein from nearly half a million patients screened for CF definitively shows that sensitivity (i.e. coverage) and specificity are the most critical factors in selecting an effective panel for targeted *CFTR* testing[1,7]. This counters the notion that inclusion of more variants on a given panel equates to better coverage, as the AmplideX PCR/CE *CFTR* Kit* incorporates variants that translate to >5% improvement in individual carrier detection from diverse populations compared to other panels, including several with more total variants. Since significant treatment advancements for CF have the potential to significantly improve outcomes, delayed diagnosis and missed carriers can lead to serious clinical consequences. In order to reduce the disparate outcomes observed in non-white CF patients[12,13], it is imperative to choose a test panel that reflects the *CFTR* variant heterogeneity of diverse populations.

REFERENCES:

1. Beauchamp, K.A., et al., *Sequencing as a first-line methodology for cystic fibrosis carrier screening.* Genet Med, 2019. **21**(11): p. 2569-2576.
2. Farrell, P.M., et al., *Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report.* J Pediatr, 2008. **153**(2): p. S4-S14.
3. Schrijver, I., et al., *The Spectrum of CFTR Variants in Nonwhite Cystic Fibrosis Patients: Implications for Molecular Diagnostic Testing.* J Mol Diagn, 2016. **18**(1): p. 39-50.
4. Castellani, C., et al., *Benchmarks for cystic fibrosis carrier screening: a European consensus document.* J Cyst Fibros, 2010. **9**(3): p. 165-78.
5. Deignan, J.L., et al., *CFTR variant testing: a technical standard of the American College of Medical Genetics and Genomics (ACMG).* Genet Med, 2020.
6. Tezacaftor/ivacaftor for cystic fibrosis. Aust Prescr, 2019. 42(5): p. 174-175.
7 Westemeyer, M., et al., *Clinical experience with carrier screening in a general population: support for a comprehensive pan-ethnic approach.* Genet Med, 2020.
8. Castellani, C., et al., *ECFS best practice guidelines: the 2018 revision.* J Cyst Fibros, 2018. **17**(2): p. 153-178.
9. Committee on, G., *Committee Opinion No. 691: Carrier Screening for Genetic Conditions.* Obstet Gynecol, 2017. **129**(3): p. e41-e55.
10. Accurso, F.J., M.K. Sontag, and J.S. Wagener, *Complications associated with symptomatic diagnosis in infants with cystic fibrosis.* J Pediatr, 2005. **147**(3 Suppl): p. S37-41.
11. Campbell, P.W., 3rd and T.B. White, *Newborn screening for cystic fibrosis: an opportunity to improve care and outcomes.* J Pediatr, 2005. **147**(3 Suppl): p. S2-5.
12. Spencer, D.A., M. Venkataraman, and P.H. Weller, *Delayed diagnosis of cystic fibrosis in children from ethnic minorities.* Lancet, 1993. **342**(8865): p. 238.
13. Spencer, D.A., et al., *Cystic fibrosis in children from ethnic minorities in the West Midlands.* Respir Med, 1994. **88**(9): p. 671-5.

*Product in development. Specifications not finalized.
†For research use only. Not for use in diagnostic procedures.
‡CE-IVD for US export only.

ASURAGEN, INC. 2150 Woodward Street, Suite 100 • Austin TX 78744 • tel 512 681 5200 • fax 512 681 5201 • asuragen.com

2000-0153